A biogeographic view on Southeast Asia's history

D i s s e r t a t i o n
zur Erlangung des akademischen Grades
d o c t o r r e r u m n a t u r a l i u m
(Dr. rer. nat.)
im Fach Biologie
eingereicht an der

Lebenswissenschaftlichen Fakultät
der Humboldt-Universität zu Berlin
von

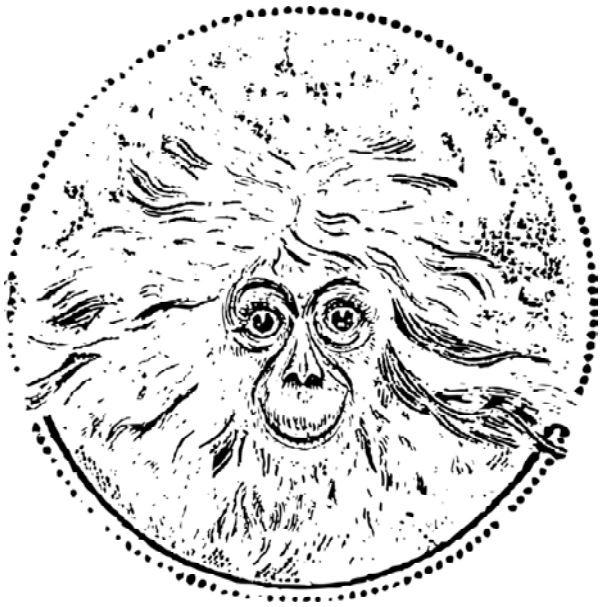
Dipl.-Biol. Björn Stelbrink

Präsident der Humboldt-Universität zu Berlin
Prof. Dr. Jan-Hendrik Olbertz

Dekan der Lebenswissenschaftlichen Fakultät
Prof. Dr. Richard Lucius

Gutachter/innen:
1. PD. Dr. Frieder Mayer
2. Prof. Dr. Johannes Müller
3. Prof. Dr. Thomas Wilke

Tag der mündlichen Prüfung: 19. 12. 2014



The above picture shows a stylized orang-utan, modified from the book cover of: Wallace A.R. (1869) *The Malay Archipelago, the land of the orang-utan and the bird of paradise a narrative of travel with studies of man and nature*. Macmillan, London.

Content

Zusammenfassung

Das tropische Südostasien, und besonders der Indo-Australische Archipel, ist bekannt für seine bemerkenswerte floristische und faunistische Diversität, besonders konzentriert in vier der identifizierten Biodiversitäts-Hotspots (Indochina, Sundaland, die Philippinen und Wallacea). In dieser Arbeit wird die biogeographische Geschichte Südostasiens beleuchtet, um Regionen mit einer erhöhten Biodiversität zu identifizieren und zu testen, ob dies mit Diversifikationen innerhalb der Region und Einwanderungen und/oder Auswanderungen korreliert und ob sich diese Faktoren über die Zeit hinweg ausgleichen. Ein besonderer Augenmerk wird auf Sulawesi und seine besondere Fauna gelegt, um zu testen, ob ein Ursprung durch Vikarianz für verschiedene Tiergruppen plausibel erscheint und wann Diversifikationen innerhalb der Fisch- und Schnecken-Radiationen im Malili-Seensystem begannen. Dabei wird auf Meta-Analysen und mehrere Disziplinen zurückgegriffen für eine integrative biogeographische Geschichte Südoastasiens und seiner Fauna, indem molekulare Uhr-Analysen, Berechnungen zur Ermittlung des Ursprungsortes mit tektonischen, paläogeographischen und klimatischen Rekonstruktionen verbunden werden, um potentielle Ursachen für die heutige Verbreitung zu finden.

Borneo und Indochina sind große “evolutionäre Hotspots” für verschiedene Tier- und Pflanzengruppen, hauptsächlich verursacht durch Diversifikationen innerhalb des Gebietes und durch anschließende Auswanderungsergnisse seit dem frühen Miozän, und noch verstärkt im Plio-Pleistozän. Im Gegensatz dazu sind die restlichen Gebiete eher durch Einwanderungen und Diversifikationen charakterisiert, vor allem bei vergleichsweise jungen, vulkanisch-aktiven, Inseln wie Java und Sumatra. Sulawesi ist besonders interessant um biogeographische Hypothesen zu testen, denn es liegt zwischen zwei großen zoogeographischen Regionen und hat eine sehr komplexe tektonische und paläogeographische Geschichte erfahren. Dazu gehören die Öffnung der Straße von Makassar mit der damit verbundenen Abspaltung West-Sulawesis von Borneo (Sunda-Shelf), die Verlängerung und westliche Wanderung des Sula-Spurs und die damit einhergehende Abspaltung vom australischen Kontinentalrand (Sahul-Shelf) sowie die abschließende Verschmelzung des heutigen Sulawesis im Miozän. Dieses geologische Szenario eröffnet die Möglichkeit, einen Ursprung durch Vikarianz für verschiedene Gruppen zu testen. Allerdings datieren die meisten berechneten Aufspaltungsereignisse alle betrachteten Vikarianz-Ereignisse und deuten deshalb darauf hin, dass die Insel hauptsächlich durch Dispersal besiedelt wurde, vermutlich verursacht und verstärkt durch periodische Meeresspiegel-Schwankungen im Pliozän. Interessanterweise konnte eine „tectonic dispersal” vicariance-Hypothese ausgehend vom australischen Kontintenalrand für drei Taxa nicht ausgeschlossen werden.

Das Malili-Seensystem (hauptsächlich der Matano-See, der Mahalona-See und der Towuti-See) in Zentral-Sulawesi ist ein Süßwasser-Biodiversitäts-Hotspot, gekennzeichnet

durch seine endemischen Artenschwärme wie die Schnecken-, Fisch-, Garnelen- und Krabbenradiationen. Die Fische Sulawesis (Telmatherinidae) entstammen wahrscheinlich der Sahul-Shelf-Region und haben die Insel vermutlich über Dispersal und "tectonic dispersal" erreicht und von dort die Seen von den Flüssen aus besiedelt. Diversifikationen innerhalb der epibenthischen Spitzflosser und der hauptsächlich pelagischen und benthopelagischen Rundflosser im Matano-See sind älter als beobachtete Hybridisierungen mit Fluss-Populationen. Damit kann davon ausgegangen werden, dass es nach der Besiedlung des Sees zu einer bethisch-pelagischen Aufspaltung gekommen ist, gefolgt von schnellen Radiationen innerhalb dieser zwei Gruppen. Sekundäre Hybridisierungen zwischen Spitzflossern und Fluss-Populationen haben erst später stattgefunden und haben deshalb wohl nicht zur ersten Aufspaltung der benthischen Gruppe beigetragen. Ein ähnliches Besiedlungsszenario, ausgehend vom australischen Kontinentalrand ("out-of-Australia"-Hypothese), wird für die endemische Süßwasserschnecke *Tylomelania* angenommen. Die Banggai-Sula-Region des Sula-Spurs gilt als mögliche Quelle für die Besiedlung Sulawesis durch den Vorfahren von *Tylomelania*. Datierte molekulare Phylogenien und Netzwerk-Analysen zeigen eine schnelle Aufspaltung in acht große Linien und eine starke, räumlich bedingte, genetische Struktur. Dies ist vermutlich auf das schwache Verbreitungspotential, maximale Höhengrenzen der Verbreitung und durch Gebirgsbildungen im späten Miozän und Pliozän zurückzuführen.

Abstract

Tropical Southeast Asia, and particularly the Indo-Australian Archipelago, is known for its tremendous floral and faunal biodiversity, mainly accumulated in four of the world's biodiversity hotspots identified (Indochina, Sundaland, the Philippines, and Wallacea). Here, Southeast Asia's biogeographic history is examined to identify areas being characterized by high levels of biodiversity (number of lineages, species richness) through time and to test whether the respective biota is mainly due to *in situ* diversification, immigration and/or emigration, or equilibrium dynamics. Moreover, this thesis focuses particularly on Sulawesi and its peculiar fauna to test if a vicariant origin appears plausible for certain groups and when the remarkable fish and snail radiations found in the Malili Lakes system started to diversify. To achieve this, meta-analytical and multi-disciplinary approaches are considered for an integrative historical biogeographic history of Southeast Asia and its biota by using molecular clock analyses and ancestral area estimations together with tectonic, palaeogeographic and climatic reconstructions to reveal potential causes for present-day distribution.

Borneo and Indochina are identified as major 'evolutionary hotspots' for a diverse range of fauna and flora mainly caused by intra-island diversification and subsequent emigration since at least the early Miocene and more pronounced in the Plio-Pleistocene. In contrast, the remaining areas are rather characterized by both the accumulation of immigrants and *in situ* diversification, most evident for comparatively young volcanically active emergent islands such as Java and Sumatra. Sulawesi is of particular importance for testing biogeographic hypotheses as it is located between two major zoogeographic regions and has experienced a highly complex tectonic and palaeogeographic history. This includes the opening of the Makassar Strait and the consequent separation of West Sulawesi from Borneo (Sunda Shelf), the extension and westward movement of the Sula Spur and thus the separation from the Australian margin (Sahul Shelf), and the final fusion of present-day Sulawesi in the Miocene. This geological scenario allows testing for a vicariant origin of certain Sulawesi groups; however, most divergence time estimates postdate relevant vicariant events and hence suggest that the island was predominantly colonized by dispersal potentially triggered and intensified by periodic sea-level fluctuations in the Pliocene. Interestingly, a "tectonic dispersal" vicariance hypothesis from the Australian margin could not be refuted for three taxa.

The Malili Lakes system (mainly Lake Matano, Lake Mahalona and Lake Towuti) in central Sulawesi is a hotspot of freshwater biodiversity, characterized by endemic species flocks such as the pachychilid snail, telmatherinid fish, atyid shrimp and gecarcinucid crab radiations. Sulawesi's Telmatherinidae have most likely originated in the Sahul Shelf area, have probably reached the island by both marine dispersal and terrane-rafting ("tectonic dispersal"), and have colonized the lakes from rivers. Diversification within epibenthic sharpfins and the

predominantly pelagic to benthopelagic roundfins in Lake Matano clearly predates hybridization events with stream populations. For Lake Matano, a scenario of initial benthic-pelagic divergence after colonization of the lake by riverine populations can be assumed, followed by rapid radiation within both clades. Secondary hybridization of stream populations with sharpfins occurred more recently and hence most likely has not contributed to the initial divergence of this benthic species flock. Similarly, for the endemic freshwater snail *Tylomelania* a vicariant origin from the Australian margin ("out-of-Australia" hypothesis) is proposed. The Banggai-Sula region of the Sula Spur is identified as a possible source area for the colonization of Sulawesi by proto-*Tylomelania*. Time-calibrated molecular phylogenies and parsimony network analyses show a rapid diversification of *Tylomelania* into eight major lineages and a strong spatial genetic structure. This is likely related to poor dispersal capabilities and altitudinal distribution limits, and prominent mountain uplifts in the late Miocene and Pliocene.

List of Figures

List of Tables

General Introduction

"A reasonable biogeographer is neither a vicarianist nor a dispersalist but an eclecticist"
Simpson (1980)

Major questions in evolutionary biology include, amongst others, the origin and distribution of the world's biodiversity: How many species exist and how do we explain the current species richness in the world and its disparity among taxa? Why are some species widespread, while others are geographically restricted to a comparatively small area? What are the underlying factors and processes in distribution, diversification and speciation observed in a certain taxonomic group? Here, the field of (historical) biogeography investigates the biotic evolution in a spatial and – at its best – in a temporal context by combing phylogenetic, geological, palaeontological and climatic information of a given study area, and it thus has become a comprehensive research field particularly in recent years.

Present-day biogeographic research has been greatly influenced by Alfred Russel Wallace and his seminal work *The geographical distribution of animals* (Wallace, 1876). This book is predominantly based on early expeditions to the Malay or Indo-Australian Archipelago (IAA) between 1854 and 1862 (Wallace, 1859, 1860, 1863, 1869), during which Wallace collected more than 125,000 specimens of insects, molluscs, reptiles, birds, and mammals (Wallace, 1869). The analysis of the archipelago's faunal composition prompted Wallace to conclude "that the Strait of Lombock (only 15 miles wide) marks the limits and abruptly separates two of the great Zoological regions of the globe" (Wallace, 1860), namely the Indian or Oriental Region and the Australian Region.

'Wallace's line' – thus dubbed for the first time by Thomas H. Huxley (Huxley, 1868) – originally ran through the Lombok Strait (between Bali–Lombok), along the Makassar Strait (between Borneo–Sulawesi), and ended south of the Philippine island of Mindanao in the Celebes Sea (Fig. 0.1; Wallace, 1876). However, later naturalists extended and modified this zoogeographic boundary (mainly based on different interpretations of bird and mammal distribution patterns), excluding Bali and the Philippines (except Palawan) from the Oriental Region or by moving this line of 'faunal balance' to the far east of the IAA (see Fig. 0.1 and Mayr, 1944 and Simpson, 1977 for details). Interestingly, later studies of Pleistocene sea-level stands in that region have shown that these zoogeographic lines (Wallace's Line to some extent, Huxley's modified Line and Lydekker's Line; Lydekker, 1896) nicely follow the continental margins of both the Sunda (Asian) Shelf in the west and the Sahul (Australian) Shelf in the east; and "between the two last-described shelves lies Wallacea […], a transition zone wherein Asiatic and Australian types mingle" (Dickerson, 1928).

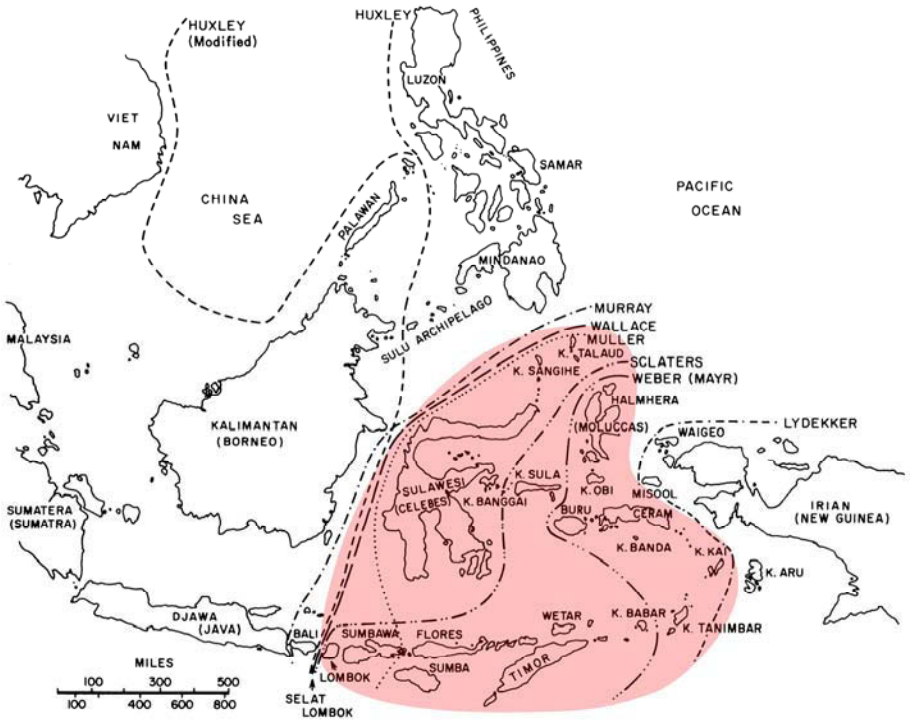


Figure 0.1. Overview of zoogeographic boundaries across the Indo-Australian Archipelago proposed by different authors (Wallacea is marked in red; note that Dickerson previously included the Philippines except for Palawan). Modified from Simpson (1977); see Mayr (1944) and Simpson (1977) for details.

The peculiar assembly of floral and faunal elements in the IAA can be associated with Southeast Asia's complex geological history forming a heterogeneous archipelago of various types of islands. According to Whittaker & Fernández-Palacios (2007), 'true islands' (i.e., "land being wholly surrounded by water") can be classified based on their mode of origin (Table 0.1). 'Continental islands' represent isolated fragments of the continental shelf being isolated from other such fragments by shallow water straits. The biota are therefore mainly of continental origin, and vicariant (and dispersal) processes have likely played a key role in the biota's assembly when sea-level fluctuations (see Voris, 2000 and Sathiamurthy & Voris, 2006 for Pleistocene reconstructions in 'Sundaland', i.e., previously submerged section of the Sunda Shelf being exposed during glacials) either separated parent populations or facilitated inter-island colonization across continental islands such as Borneo, Sumatra and Java (see e.g., Inger & Voris, 2001; Mercer & Roth, 2003; Gorog et al., 2004; Hertwig et al., 2013; Klaus et al., 2013).

In contrast, 'oceanic islands' arise from volcanic activity and have never been connected to the continental shelf. Present-day flora and fauna show high levels of endemism and are predominantly founded by long-distance dispersal events. These islands represent ideal candidates for applying molecular clock analyses and consequently testing biogeographic hypotheses as both island formation and development of suitable habitats constrain time of

Table 0.1. Island classification by origin and brief information on physical and biotic characteristics. Modified from Whittaker & Fernández-Palacios (2007).

	Continental islands	**Oceanic islands**	**Continental fragments**
Origin	Sea-level rise	Volcanic activity	Rift separation
Isolation	Low	Usually high	Variable
Longevity	Short	Variable (up to 20 My)	Variable (up to 150 My)
Water depth	<150 m	>1,000 m	>200 m (1,000 m)
Original biota	Present	Absent	Variable
Fauna	Continental	Unbalanced	Unbalanced
Endemism	Variable	High	(Very) high
Origin of taxa	Vicariance	Dispersal	Variable
Examples	Borneo, Britain	Ternate (Moluccas), Hawaii	Sulawesi, New Zealand

colonization (from the mainland or surrounding islands) and intra-island expansion (see e.g., Roderick & Gillespie, 1998; Price & Clague, 2002; Cowie & Holland, 2008; Parent et al., 2008 and Ree & Smith, 2008 for reconstructed dispersal routes across the Hawaiian Archipelago).

Finally, 'continental fragments' (micro-continents) are fragments of continental crust separated from their former landmass by tectonic drift. These micro-continents may have experienced a strikingly different geological history and consequently the biota's origin is similarly variable. This is well exemplified by Sulawesi – located in the centre of Wallacea –, with its complex geological history involving terrane movements, subsistence and emergence, and sea-level fluctuations, and which has thus accumulated a peculiar assembly of plants and animals with both Asian and Australian affinities. Given this heterogeneous nature, the IAA represents an ideal candidate region for addressing questions related to island biology (see e.g., Lohman et al., 2011), such as the biogeographic origin of an island's biota (vicariance *vs.* dispersal scenarios; Asian *vs.* Australian origin), species-island age relationships, probabilities of inter-island and island-continent dispersal events, and 'intra-island dynamics' such as diversification, expansion, persistence and extinction of lineages (Heaney, 2007).

Transition zones, zoogeographic realms, (freshwater) ecoregions, 'areas of endemism', and other such biogeographic subdivisions as briefly mentioned above have been the subject of many debates, primarily arising from various descriptions of local, regional and global biodiversity patterns at different taxonomic scales and using a range of methodological approaches (see e.g., Mayr, 1944 and Simpson, 1977 for reviews and e.g., Jepson & Whittaker, 2002; Wikramanayake et al., 2002; Parenti & Ebach, 2010, 2013; de Bruyn et al., 2013b; Holt et al., 2013; Kreft & Jetz, 2013 and Waters et al., 2013 for recent discussions). However, understanding the biotic evolution in a particular region (or single island) not only requires the description of faunal and floral patterns but, more importantly, the identification of related processes (event-based approaches).

Morrone & Crisci (1995) identified and described five historical (phylogeny-based) biogeographic concepts in detail, representing the first attempts to integrate observed distribution patterns into a cladistic framework, some of which are still feature quite commonly

in recent biogeographic literature: dispersalism, phylogenetic biogeography, panbiogeography, cladistic biogeography, and parsimony analysis of endemism (e.g., Turner et al., 2001; Heads, 2002, 2012; van Welzen et al., 2003; Ebach & Morrone, 2005; Morrone, 2005; Parenti & Ebach, 2010). Because these cladistic approaches fail to incorporate the temporal dimension of a given phylogenetic tree (i.e., divergence time estimates; but see Folinsbee & Evans, 2012), relating those diversifications with their potential geological, climatic, or biological causes is difficult (e.g., Hunn & Upchurch, 2001; Crisp et al., 2011). Furthermore, the comparison of so-called 'general area cladograms' (phylogenetic trees corresponding to a certain region) can be misleading if similar area-specific clades have diversified at different times ('pseudo-congruence'; Donoghue & Moore, 2003).

The development of standardized and advanced methods in molecular biology (e.g., PCR amplification, DNA barcoding, *Next Generation Sequencing*) and phylogenetic analyses have resulted in a significant increase of studies dealing with intra- and interspecific relationships across a variety of taxonomic groups and geographic areas (Parr et al., 2012). More recently, advanced DNA isolation and sequencing methods of so-called 'ancient DNA' (i.e., fragmented DNA of both subfossil and dry or ethanol-fixed museum material) now enable the placement of extinct or rare taxa in a broader phylogeny and also provide new insights into biogeographic (and even domestication) history (e.g., Shapiro et al., 2002; Baker et al., 2005; Bunce et al., 2005; Larson et al., 2007; Gilbert et al., 2008; Rohland et al., 2010; de Bruyn et al., 2011; Mitchell et al., 2014; Thomson et al., 2014). Along with these methodological breakthroughs, the implementation of molecular clock analyses (see Zuckerkandl, 1987 for a historical review) has played a pivotal role for the further development of historical (molecular) biogeography. Molecular clock analyses make use of the fact that changes in nucleotide sequences (mutations) are accumulated and fixed in a population over time. Hence, the degree of relatedness and time of diversification events between two (or more) genetically isolated groups can be (roughly) estimated when genetic distance and rate of nucleotide changes (e.g., substitutions per site per year) are known (see e.g., Ayala, 1999; Bromham & Penny, 2003; Pulquério & Nichols, 2007).

Calibrated phylogenies permit testing of the congruence of diversification and speciation events with specific geological and climatic events suspected of driving biotic evolution, reconstructing geographic modes of speciation, and how rapidly certain traits evolved. Conversely, time-calibrated trees can be useful when testing for proposed geological histories of a particular area as frequently debated for New Zealand and New Caledonia (e.g., Knapp et al., 2007; Grandcolas et al., 2008; Giribet & Boyer, 2010; Nattier et al., 2011; Swenson et al., 2014). Despite the improvements in Bayesian molecular clock analyses (e.g., strict *vs.* relaxed [local] clock, Yule *vs.* birth-death process, demographic models, etc.; Drummond et al., 2005, 2006; Drummond & Suchard, 2010; Heled & Drummond, 2010), the

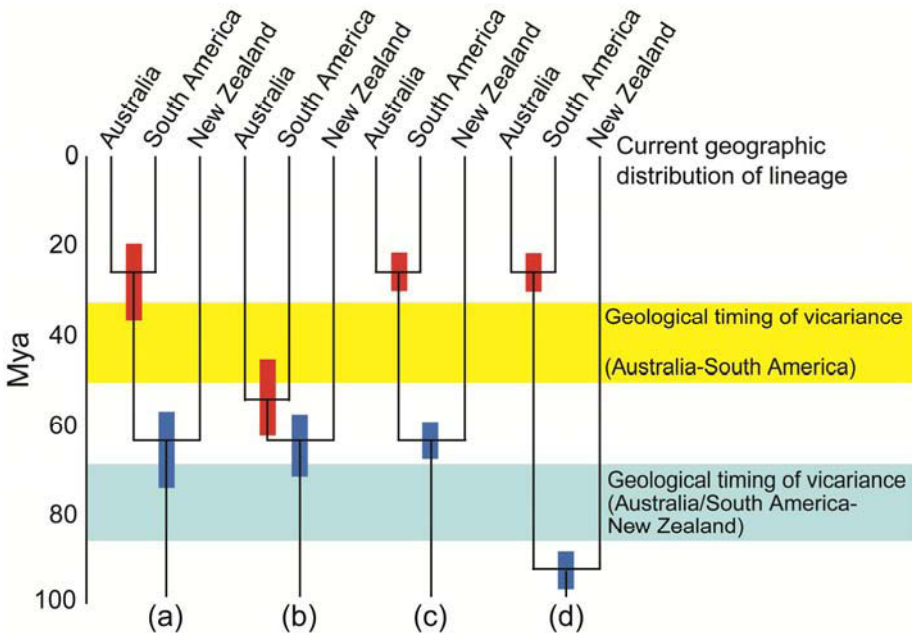


Figure 0.2. Illustration of how vicariance scenarios can be tested using divergence time estimates and hypothesized timing of geological events, here the separation of New Zealand from Australia/South America c. 70-85 Mya and the separation of Australia from South America c. 30-50 Mya. Vicariant scenarios cannot be rejected if divergence-time error bars (red and blue vertical bars) overlap with particular geological events (i.e., rift separation; yellow and light blue horizontal bars). Redrawn from Crisp et al. (2011).

estimation of divergence times strongly depends on comprehensive taxonomic and geographic sampling, the choice of appropriate (taxon- and marker-specific) substitution rates, and on a broad chronological range of geological/fossil calibration points. While incomplete taxonomic and geographic coverage might conceal certain interspecific relationships in a given phylogeny, externally applied rates and dates directly influence internal node ages, and thus may yield misinterpretations. Fossil-based calibrations only provide a minimum age for a specific divergence event (e.g., Yang & Rannala, 2006; Ho & Philipps, 2009), while the use of geological (vicariant) events as calibration points such as the closing of the Isthmus of Panama, the separation of Madagascar from the African continent, or the continuous break-up of Gondwana can lead to circular reasoning (e.g., Waters & Craw, 2006; Rutschmann et al., 2007).

Two major biogeographic hypotheses can be taken into account for explaining observed disjunct distribution patterns among two sister groups, vicariance and 'long-distance dispersal and establishment' (or trans-oceanic dispersal; see also Crisp et al., 2011). A vicariance scenario assumes the continuous (geographic) separation of ancestral populations caused by a physical barrier (mountain chains, water straits, landmass fragmentation such as the Gondwanan break-up, see a review by Upchurch, 2008) and a subsequent genetic isolation of resulting subpopulations when gene flow is disrupted. Hence, these geological events represent constraints for a vicariant origin and thus can be used to test this particular hypothesis when time estimates for both geological and diversification events are available. The hypothesis of a vicariant origin for a given diversification event is rejected if the credibility interval

(divergence-time error bar) of the respective node post-dates the estimated time range of the vicariant (geological) event under consideration; however, this hypothesis cannot be ruled out if those credibility intervals overlap with a particular vicariant event (Fig. 0.2).

In contrast, under the hypothesis of a dispersal scenario, (single or few) individuals of an ancestral population are dispersed (e.g., by wind or rafts) and thus become geographically, and eventually genetically, isolated from parent populations. While the constraints of geological events can be used to refute a vicariant scenario, hypothesis testing appears *per se* impossible for dispersal. However, Crisp et al. (2011) suggest a few examples, potentially providing a framework of hypothesis testing for (long-distance) dispersal: 1) ocean and wind currents may strongly bias dispersal events in particular regions, here the Southern Hemisphere; 2) stepping-stone dispersal events appear more likely between adjacent islands/landmasses and could be biased in a direction if respective stepping stones are of different geological age (see e.g., the Hawaiian ridge; Roderick & Gillespie, 1998; Cowie & Holland, 2006); and 3) ecological parameters can be included to estimate the different probabilities of dispersal events at a certain point of time. A more detailed description is presented by Gillespie et al. (2012), identifying modes and directionality of colonization routes, and characterizing post-arrival diversification processes when ocean currents, prevailing winds and vectors such as sea birds are considered.

Historical biogeography has thus recently become a multi-disciplinary field by combining information on tectonics and palaeogeography (mainly orogenic events and distribution of land and sea), palaeontology (fossil remains, pollen record), and ecological parameters (e.g., climate-mediated temperature and sea-level fluctuations) with time-calibrated phylogenies and a variety of phylogenetic models currently available (Ronquist & Sanmartín, 2011; Matzke, 2013) to test for the hypothesized scenarios. Under this framework, several widespread and diverse groups have been recently studied and reviewed: e.g., Southern Hemisphere lungfishes, ratite birds and cichlid fishes and the relation to the Gondwanan break-up (Upchurch, 2008; Friedman et al., 2013), the endemic Malagasy fauna and its relation to the African continent (e.g., Crottini et al., 2012; Samonds et al., 2012, 2013), the origin of corvoid birds in New Guinea and their distribution across Southeast Asia (Jønsson et al., 2011), and the origin of Hawaiian *Drosophila* flies and honeycreepers and their inferred dispersal routes across the island chain (Lerner et al., 2011; Obbard et al., 2012).

These studies nicely demonstrate that the application of molecular clock analyses is most useful in a multi-disciplinary approach if the biogeographic history of a certain taxonomic group is studied. This is even more essential when considering biotic changes in a particular region and thus general patterns (i.e., divergence time estimates and modes of colonization) identified among a wide range of taxonomic groups analyzed (i.e., meta-analyses). This is of special interest for so-called 'biodiversity hotspots', i.e. areas showing high levels of species richness and endemicity and being under threat from habitat degradation and loss. In this

Figure 0.3. Biodiversity hotspots identified by Myers et al. (2000) based on selected vascular plant and vertebrate (amphibians, reptiles, birds and mammals) species richness and endemicity. Note that freshwater diversity is not considered. Redrawn from Myers et al. (2000).

context, Myers et al. (2000) identified 25 hotspots based on vascular plants and some vertebrate groups (see Fig. 0.3). While this list of biodiversity hotspots is most likely biased considering the lack of speciose groups such as arthropods and molluscs (e.g., May, 1990; Ødegaard, 2000), information on species richness, endemicity and ecological parameters, however, might be potentially useful for setting up management frameworks and studying biogeographic patterns in particular regions of interest (Wiens & Donoghue, 2004).

In this thesis, the biogeographic history of Southeast Asia is studied at both large and small scales using both meta-analytical approaches and case studies to identify general and specific (geological and biological) processes accounting for distribution patterns observed across the IAA. Inferred phylogenetic relationships and colonization routes among a variety of taxonomic groups are tested against palaeogeographic and climatic hypotheses and *vice versa*.

The first chapter represents a large-scale meta-analysis comprising a mixed set of 61 plant and animal datasets distributed in the IAA and adjacent regions (i.e., plants, invertebrates, and vertebrates; terrestrial, freshwater, volant, and non-volant taxa). Origin, accumulation, *in situ* (intra-island) diversification and subsequent emigration of lineages across the archipelago is explored to identify major sources of biodiversity in the light of geological and climatic reconstructions.

The second chapter represents a meta-analysis as well, but with a special focus on Sulawesi's fauna, which is believed to be mainly of Sundaic (Asian) origin. Time-calibrated phylogenetic analyses are applied to test for zoogeographic affinities (Sunda Shelf *vs.* Sahul Shelf) and modes of colonization (vicariance *vs.* dispersal). Sulawesi is considered a biogeographic melting pot, potentially being a 'transition zone' between two zoogeographic regions (see above), thus highlighting the importance of studying its biotic origins.

In the third chapter, different calibration strategies (fossils and geological events) are tested and evaluated to date both speciation and hybridization events found in the radiation of endemic sailfin silversides inhabiting the Malili Lakes system of Sulawesi. Furthermore, phylogeographic relationships between these fishes and their sister group found in Sulawesi and off New Guinea are inferred by testing divergence time estimates against proposed geological reconstructions. Freshwater taxa, particularly those lacking brackish or saltwater tolerant larval stages, are of great interest for testing different colonization scenarios given their dispersal constraints.

The fourth and final chapter examines Sulawesi's enigmatic freshwater snails of the genus *Tylomelania* and tests the previously proposed 'Out-of-Australia' hypothesis in a temporal (molecular clock) and palaeogeographic framework. Geological processes giving rise to Sulawesi's colonization (mainly movements and fusions of terranes) and putatively triggered intra-island diversification, expansion and geographical isolation since the late Miocene and Pliocene (particularly mountain uplifts) are evaluated.

This again stresses the need for an integrative approach including different research disciplines and combining meta-analyses with single-taxon studies to understand present-day biodiversity, past colonization routes and dispersal barriers, geological and evolutionary processes through space and time and hence to obtain a more complete picture of a region's or biota's historical biogeographic history.

Chapter 1

Borneo and Indochina are major evolutionary hotspot for Southeast Asian biodiversity

Introduction

Since the early research (1859-1871) of Alfred Russel Wallace, Southeast (SE) Asia has held considerable interest for the study of evolutionary biology and the field of biogeography. Megadiverse SE Asia comprises both mainland and numerous archipelagos of oceanic and continental islands (Lohman et al., 2011). The fragmented nature of SE Asia's geography is mirrored in the distribution of its extant biota and a high degree of regional and local endemism (e.g., Woodruff, 2010; Wong, 2011). The four biodiversity hotspots that constitute SE Asia (Indo-Burma, Sundaland, Wallacea and the Philippines; Myers et al., 2000; Woodruff, 2010) are under increasing threat from the unsustainable use of native species and their habitats (Koh & Sodhi, 2010; Wilcove et al., 2013).

Biotic evolution in SE Asia cannot be understood without considering the region's complex tectonic and climatic evolution (Fig. 1.1; Fig. S1-S4, Supplementary Information), dominated since the late Paleozoic by a series of ongoing continental collisions (Metcalfe, 2011a). These geological events have influenced many environmental variables e.g., geomorphology, topography, atmospheric circulation, the hydrological cycle and ocean current patterns, the opening and closure of ocean gateways, and as a consequence, climate and climate change (see Fig. 1.2 for Quaternary sea-level scenarios), either directly or indirectly (Hall, 2009b; Morley, 2012). Present-day SE Asia is the result of more than 300 million years of 'Colliding Worlds' (van Oosterzee, 1997) characterized by continent-continent and continent-arc collisions. SE Asian continental blocks (Fig. S1-S3) were all derived from the Southern Hemisphere supercontinent Gondwana, and travelled north to progressively collide and coalesce prior to the ongoing collision with the northwards-moving Australian continent (Metcalfe, 2011a). By the end of the Triassic (c. 200 Mya), core Sundaland, comprising Sumatra, the Thai-Malay Peninsula and most of the present-day Sunda Shelf, was part of continental Asia (Fig. S2 and S3). Vegetation during the early Cretaceous was dominated by the extinct gymnosperm family Cheirolepidiaceae, with thick cuticles suggesting a palaeoclimate with some degree of water stress (Vakhrameev, 1991). Angiosperms increased in number and diversity during the Late Cretaceous (Morley, 2000). Occurrence of *Nypa* pollen testifies to the presence of mangroves, while common Gnetales and Laurasian conifer pollen suggest upland open woodland vegetation also subject to water stress.

Three major Cenozoic collision events had a great influence on the current archipelago setting and probably substantially affected both climate and the assembly of regional biota, possibly in tandem. The impact of the Indian-Asian collision during the Eocene (Fig. S4) on SE Asia's geography is debated but likely to have been small (Hall, 2009b). The terrestrial connection between India and Asia, however, had a huge effect on some SE Asian fauna (e.g., Shih et al., 2009; Klaus et al., 2010; Li et al., 2013) and flora (Morley, 2000), which changed

dramatically in the earliest middle Eocene (c. 49 Mya) compared to that of the Paleocene and late Mesozoic. Dispersal of many plant taxa of Indian origin into SE Asia resulted in extinction of many elements of the older Paleocene flora (Fig. S4) and by 45 Mya, pollen data indicate a diverse new flora (Morley, 2000). Following the late Eocene change to a global 'Ice House' climate, early Oligocene pollen floras were of markedly lower diversity but by the end of the Oligocene most pollen types characteristic of the wet, rainforest-dominated early Miocene were in place (Supplementary Information).

The vegetation of SE Asia was similarly influenced by the Asia-Australia collision (about 25-23 Mya), although the extensive forests of the Sunda Shelf islands were not substantially invaded by Australian elements (Richardson et al., 2012). However, the collision had a major impact on the climate and vegetation of the region by disrupting the Indonesian Throughflow, the major inter-oceanic current that passes through the Malay Archipelago (*sensu* Alfred Russel Wallace). Moisture that previously moved between the Pacific Warm Pool (area enclosed by the 28.5 °C sea-surface temperature isotherm) and the Indian Ocean in the Throughflow subsequently fell on Sundaland (Morley, 2006). This climate change coincided with the development of the East Asian Monsoon across China (Sun & Wang, 2005). The Australian-Sunda collision thus appears to have initiated the East Asian monsoon (Morley, 2012). This change to a wetter climate led to the development of the modern-aspect Malesian flora (Morley, 2000).

The third collision, of the westernmost Australian promontory the Sula Spur, with Sulawesi during the Miocene, led to the first amalgamation of continental fragments derived from Sundaland and Australia, respectively (Hall, 2009b, 2011). Emergence of land in central and eastern parts of Sulawesi (Hall, 2009b, 2011) is supported by biological evidence such as timing of the first major colonization wave of Sulawesi by terrestrial animal taxa from both east and west of Wallacea in the late Miocene and Plio-Pleistocene (Sahul and Sunda Shelf; Stelbrink et al., 2012). Several islands west of 'Wallace's Line' (e.g., Borneo, Sumatra, Java) are largely of continental origin, being linked to the Sunda Shelf, and formed a contiguous landmass when sea-level dropped (so-called 'Sundaland'). The Sunda Shelf attained its greatest land area during the middle Eocene (c. 49-45 Mya), with predominantly southward flowing rivers and widespread, palm-dominated (Harley & Morley, 1995) peat swamps across what is now southeastern Borneo, Java and southern Sulawesi. An easternmost rift of Sundaland resulted in the middle Eocene isolation of SW Sulawesi from mainland Borneo, forming the Makassar Strait and separating Sulawesi from Sundaland (c. 45 Mya; Hall, 2009b), creating the central section of the major biogeographic break today recognized as 'Wallace's Line'. From the later middle Eocene (c. 42 Mya) onward, southern Sundaland subsided, and by the end of the Oligocene (c. 25 Mya), much of the region now occupied by Java and southern Borneo (Kalimantan) was submerged, characterized by widespread shallow shelves with many

extensive reefs (Fig. 1.1, 25 Mya). However, central Borneo and its northwestern extension to the Asian mainland remained emergent throughout this time (Fig. 1.1, 25 & 20 Mya), enabling everwet rainforest expansion northwards c. 25 Mya (e.g., Morley, 2000; see below).

The onset of wetter conditions in the latest Oligocene/Miocene, and Plio-Pleistocene climate-driven sea-level changes (Fig. 1.2) are superimposed upon this dynamic geological framework, leading to the repeated inundation of the region's continental shelves during interglacials (Hanebuth et al., 2011). During the early Quaternary, both Sundaland and the Sahul Shelf (Australian continental shelf) underwent fundamental changes in character, as the amplitude of sea-level fluctuations increased (Zachos et al., 2001), resulting in these previously submerged shelves being exposed and flooded during glacials and interglacials, respectively (Fig. 1.2) (Voris, 2000; Hanebuth et al., 2011). This phenomenon has been best studied across Sundaland, which over the past ~2 million years effectively doubled in size during glacials, and currently exhibits its smallest geographical area for that period, with biota generally in a state of refuge (Cannon et al., 2009; Woodruff & Turner, 2009). For most of the Quaternary, sea levels fluctuated ~40-60m below present-day levels, and climate was neither so dry nor cool as during the Last Glacial Maximum (LGM). The dual impact of geological and global climatic cycles have generally been regarded as important factors contributing to the region's biotic assembly (Hall, 2009b; Lohman et al., 2011).

Building on earlier work on SE Asian biogeography (e.g., Wallace, 1869; Mayr, 1944; Simpson, 1977; Whitmore, 1981, 1987), recently refined geological and tectonic models from this region provide information on the timing and distribution of various landmasses and islands (Hall, 2009b, 2011, 2012a) (Fig. 1.1, Supplementary Information). These models underpin biological models of historical distribution of ecosystems through time (Fig. 1.2), informed by palaeontological and palynological data and based on palaeoclimatic interpretations (e.g., Cannon et al., 2009). Recent studies provide estimates of divergence times, vicariance and dispersal events and can, in combination with other empirical geological and biological data, identify key areas for the generation and maintenance of biotic diversity (e.g., Stelbrink et al., 2012). However, the complex geological history of SE Asia raises issues in the testing and interpretation of single-taxon studies, particularly across the region's heterogeneous mix of volant *vs*. non-volant terrestrial and freshwater taxa.

Thus, we here conduct meta-analyses and syntheses of geological, climatic and biotic (including 61 phylogenetic) datasets from the SE Asian region to examine the assembly of the region's ecosystems and biota through time and space. We examine whether the SE Asian biota has been characterized primarily by *in situ* diversification, immigration and/or emigration, or equilibrium dynamics between our pre-defined biogeographic areas across three major time-periods for which phylogenetic and palaeogeographic data were available, the pre-Miocene, Miocene and Plio-Pleistocene. We use divergence time and ancestral area estimation to test

whether the sources of long-term biological diversity in SE Asia accord with predictions based on our geological, climatic and biotic meta-analyses and syntheses. These predictions, based on major events in SE Asia's known history presented above, include: 1. Levels of *in situ* diversification events – i.e. the accumulation of new lineages within areas – should reflect emergent ages and sizes of our pre-defined areas; 2. Areas with recent histories of sea-level inundation (e.g., Sumatra and Java) should be characterized by higher levels of (recent) immigration than emigration; 3. Number of colonization events should be greater between adjacent areas compared to more distant areas; 4. An increase in diversification rates should be evident during the Plio-Pleistocene resulting from increased connection-disconnection events related to sea-level fluctuations.

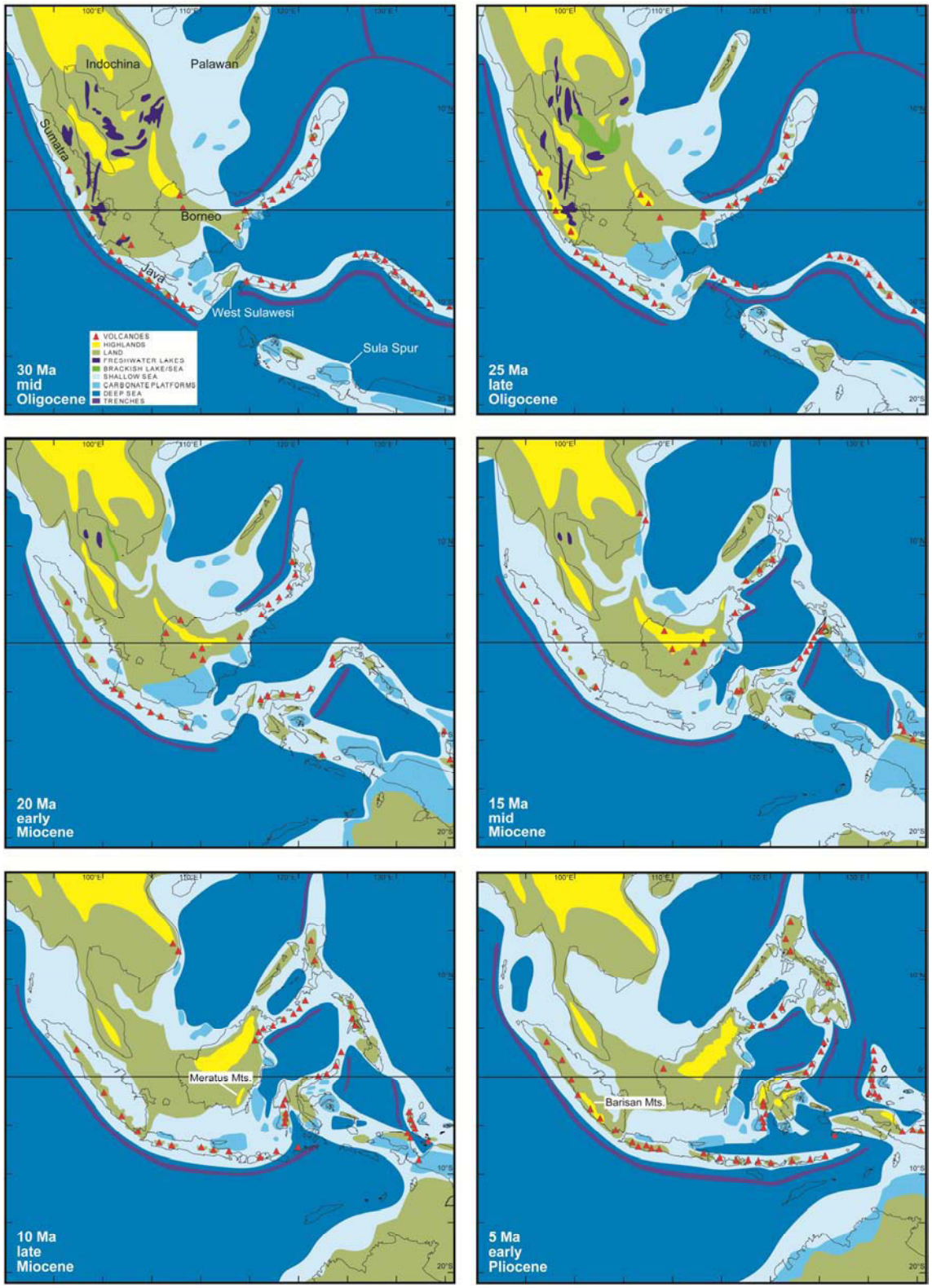


Figure 1.1. Paleogene and Neogene maps of Southeast Asian palaeogeography and key habitat availability. Palaeogeography of the region from Sundaland to northern Australia and the west Pacific between 30 Mya and 5 Mya overlain on tectonic reconstructions of Hall (2012a) in which the basis for the reconstructions is described in detail (see Materials and Methods for further information).

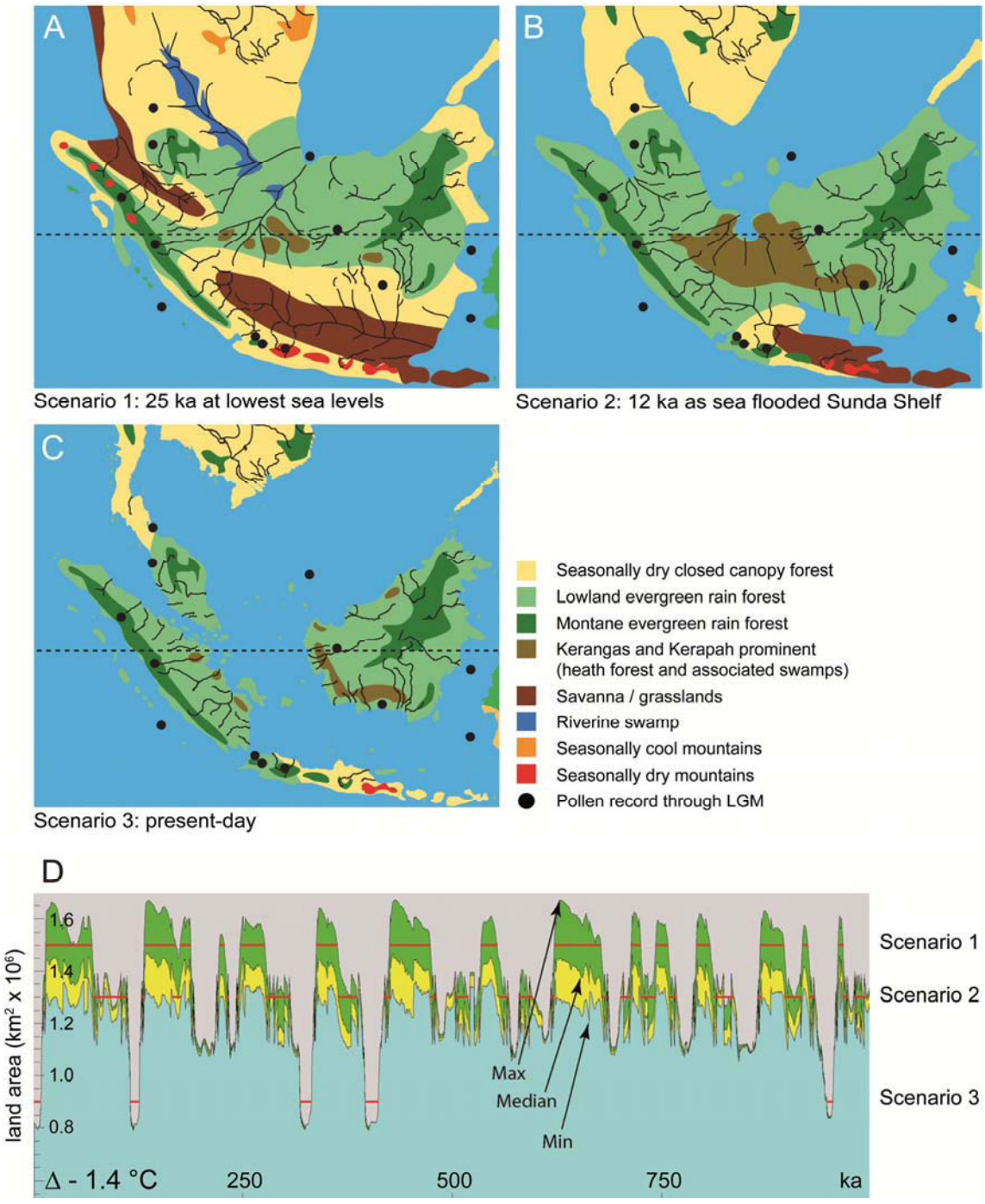


Figure 1.2. Three Quaternary climate sea-level scenarios for Southeast Asia: A) Scenario 1: 25 ka at lowest sea levels, glacial maximum, Marine Isotope Stage (MIS) 2. B) Scenario 2: 12 ka as sea flooded Sunda Shelf, intermediate sea level, MIS Stages 1, 3, 4, 5a-d. C) Scenario 3: present-day, interglacial high sea level, MIS Stages 1, 5e. Rainforest distribution based on fig. 1a in Cannon et al. (2009). Last Glacial Maximum pollen localities from Morley (2012). D) Land area of tropical evergreen broadleaf forest over last 1 million years, from Cannon et al. (2009); horizontal red lines show times when scenarios 1-3 would likely have occurred (note that the blue line represents the most conservative model but not the most likely one here). Note that Borneo, in particular, has hosted extensive rainforests throughout all periods shown.

Materials and Methods

Southeast Asian geological reconstructions

Palaeogeographic reconstructions for the Paleozoic and Mesozoic follow those published in Metcalfe (2011b). Maps were generated using a wide range of multidisciplinary data including stratigraphic, sedimentological, biostratigraphic, biogeographic, palaeomagnetic, palaeo-climatic, structural, isotopic and geochronological data and plutonic and volcanic activity (e.g., Wakita & Metcalfe, 2005). Tectonic reconstructions for the late Jurassic to present depicting the evolution of continental blocks, ocean basins and distributions of land and sea are modified versions of Spakman & Hall (2010) and Hall (2012a) and were produced using geodynamic reconstruction modeling software. The methodology and software used is described in Hall (2002, 2012a). The distribution of land and sea is shown for a larger area than in previous publications, extending further north, as discussed in Hall (2012b, 2013) and includes additional details of Sunda shelf marine embayments (Morley & Morley, 2013) and shows freshwater rift lakes and inland seas for the Sunda Shelf (Shoup et al., 2012).

Development of the SE Asian flora inferred from palynology

For the Cenozoic, patterns of palaeoclimate change have been established from palynological records generated by petroleum industry studies (partly unpublished) from across the region (Morley, 2012). These records total in excess of 150 datasets, extending from East and West Java Seas, Makassar Straits, offshore Sabah and Sarawak, Nam Con Son and Cuu Long Basins offshore Vietnam, the West Natuna Basin in Indonesia, the Malay and Penyu Basins offshore West Malaysia and Gulf of Thailand Basins, the more important of which are published (e.g., Morley & Morley, 2011). All the datasets are placed within a high resolution sequence biostratigraphic framework (following Morley et al., 2012) allowing palaeoclimates to be consistently reconstructed for individual time slices across the entire region. The high-resolution biostratigraphy used to provide the temporal framework for the Sunda Shelf palaeogeographic maps was undertaken by characterising transgressive-regressive cycles using a sequence biostratigraphic approach initially outlined in Morley (1996). The detailed stratigraphic framework used for preparing the Sunda shelf maps is presented in Morley et al. (2012). The time slices for the Sunda Shelf were mapped seismically, interpreted in terms of palaeoenvironment and presented in Shoup et al. (2012). Those maps were then simplified, rescaled and incorporated into the regional palaeogeographic maps of Hall (2011a). For Sumatra, basin outlines and stratigraphy was obtained from Moulds (1989) and Aldrich et al. (1996), with environments interpreted from unpublished data. For the Quaternary, maps have been prepared using published palynological and other climate-indicative datasets combined

with climate modeling techniques (e.g., as in Cannon et al., 2009 and biogeographic considerations, as in Morley, 2012).

Phylogenetics: Molecular clocks and ancestral area reconstructions

A literature search was performed for published molecular phylogenies including taxa (subsequently assigned to the following categories: plants, insects & spiders, freshwater crustaceans, freshwater molluscs, freshwater fishes, herpetofauna, birds, and mammals) from SE Asia (and adjacent regions). Datasets were not included *a priori* if, firstly, sample sites were restricted to less than three areas pre-defined for the ancestral area estimation, and secondly, sequences, fossil data for calibration/substitution rates, and outgroup taxa were not available for performing molecular clock analyses. Some potential biases/artifacts might occur while performing meta-analyses such as that conducted here: 1) sampling artifacts, i.e., weak taxonomic and geographic coverage due to sampling area access problems; 2) incomparable divergence times among intraspecific and interspecific datasets. However, we screened datasets carefully and omitted a substantial number (n = 68; see Supplementary Information for a list of excluded publications/datasets) prior to further analyses to reduce potential sampling bias (see Table 1.1 and 1.2 for details). Sixty-one datasets were retained for the final molecular clock analyses and ancestral area estimation, the majority (n = 45) of which were performed or re-run for the present study using taxon-specific calibration points or substitution rates from several source publications (Table 1.1).

Sequences were aligned using MAFFT (default settings; http://www.ebi.ac.uk/Tools/msa/mafft; Katoh & Toh, 2008) and corrected by eye if necessary, i.e., removing potential gap artefacts produced by MAFFT. Alignments were reduced to unique haplotypes using DAMBE v. 5.1.1 (Xia & Xie, 2001). Best-fitting substitution models were estimated using jModelTest v. 0.1.1 (Posada, 2008; 24 models, AIC; Table 1.2). Substitution rates or fossil calibration points were obtained from the publication or secondary sources (Table 1.2). Relaxed uncorrelated lognormal analyses were performed with BEAST v. 1.6.2 (GTR+G model, ngen = 20,000,000 / 40,000,000 if ESS values for several parameters did not reach the 200, log = 400 / 800, Yule process, normal distribution for calibration points) (Drummond & Rambaut, 2007). Maximum clade credibility (MCC) trees were constructed using TreeAnnotator (BEAST package; burnin = 35,001). All BEAST log files were visualized and checked in Tracer v. 1.5 (Rambaut & Drummond, 2007) for ESS values. In a few cases (n = 8), ESS values did not reach values >200 for the parameters ‘prior’ and ‘posterior’ while the remaining parameters show ESS values considerably higher than 200. In those cases, we performed re-analyses using the less complex HKY model (plus Gamma and PropInv parameters when selected by jModelTest for the best-fitting model), because low ESS values might indicate over-parameterized substitution models (see e.g., Grummer et al., 2014). Importantly, divergence time estimates did not differ

substantially between default and modified analyses for the vast majority of datasets (cf. Table 1.2 and Table S1 and S2). Distribution of relaxed-clock (BEAST) divergence times (credibility intervals and mean ages) are shown in Fig. S5-S13).

Localities of specimens were obtained either from the source publication or from web sources (e.g., GenBank) if needed and were assigned to one of 12 areas of endemism, partly modified from areas of endemism proposed by Turner et al. (2001): Indochina = Indo-Burma *sensu* Myers et al., 2000, Thai-Malay Peninsula, Sumatra, Philippines, Palawan, Borneo, Java, 'Wallacea excl. Sulawesi', Sulawesi, 'East of Wallacea', India and Sri Lanka, and Japan, or to additional areas such as 'Africa'. A presence-absence matrix was created for the ancestral area estimation conducted using Lagrange (Lagrange configurator: http://www.reelab.net/lagrange/configurator/index) (Ree et al., 2005; Ree & Smith, 2008). Analyses were performed with default (unconstrained) settings (i.e., equal dispersal rates, single time matrices) and were constrained to a maximum range size of 2 areas because for the vast majority of datasets specimens could be assigned to single pre-defined areas only. Increasing the number of ancestral ranges allows for widespread ancestors, however, by limiting this number the determination of dispersal directionalities across two areas is simplified.

Lagrange results (i.e., estimated ancestral areas for each node with the highest relative probability) were transferred to BEAST MCC tree printouts for each dataset. Changes in geographic range representing colonizations from one area to the other have been included for resolved branches (posterior probabilities >0.5), though a higher threshold would increase the reliability of performed ancestral area estimation (see Supplementary Information for BEAST tree files). BEAST mean ages plus ancestral areas of ancestral and descendent nodes were used to assign these colonization routes to the pre-selected time frames 'pre-Miocene', Miocene, and 'Plio-Pleistocene' (until the present) (Fig. 1.3 and Fig. 1.4; see Fig. S14 and S15 for group-specific colonization routes through time among taxonomic groups). Counting dispersal events within each time bin allows for comparison of time-frame specific patterns with palaeogeographic reconstructions for that period (Fig. 1.1).

Speciation and dispersal through time were explored by counting and analyzing number of lineages, *in situ* (intra-area) diversification, and emigration events in each area across all taxonomic groups (Fig. 1.4 and 1.5, Table 1.3, Table S3 and S4). The results were subsequently subjected to pairwise rank-based Mann-Whitney *U* tests using SPSS Statistics v. 17.0.0 (SPSS Inc.) to test the null hypothesis of equality between areas (asymptotic significance, two-tailed; Supplementary Information).

Table 1.1. List of datasets used for phylogenetic meta-analysis including information on taxa, generalized habitat, and molecular markers used in the original study.

Dataset[a]	Common name	Family	Genus [generalized habitat][b]	Genetic marker	Taxonomic coverage[f]	nTaxa (ingroup)[g]	BEAST/ Lagrange Performed[h]	Source Ref[i]
01	plants	Meliaceae	*Aglaia* [1]	ITS	82 spec (all genera)	42	this study	42
02	plants	Araceae	*Alocasia* [1]	cpDNA+nuDNA	71/113 species	73	*	46
03	plants	Begoniaceae	*Begonia* [1]	cpDNA	92 spec (all families)	66	#	61
04	plants	Gesneriaceae	*Cyrtandra* [1]	ITS	26/300 species [A]	30	this study	2
05	plants	Ericaceae	*Rhododendron* [1]	RPB2-d	46/300 species [B]	46	this study	20
06	plants	Rafflesiaceae	several [1]	mtDNA+cpDNA+nuDNA	80% spec (all genera)	27	#	3
07	mosquitoes	Culicidae	*Anopheles* [2]	mtDNA	population-based dataset	40	this study	16
08	butterflies	Nymphalidae	*Cethosia* [1]	mtDNA+nuDNA	all species	42	#	43
09	butterflies	Nymphalidae	*Charaxes* [1]	mtDNA+nuDNA	all species	29	#	44
10	butterflies	Pieridae	*Delias* [1]	mtDNA+nuDNA	30/44 species (Wallacea)	131	#	69
11	net-winged beetles	Lycidae	*Metriorrhynchus* [1]	mtDNA	all species	56	§	5
12	giant wood spiders	Tetragnathidae	*Nephila* [1]	COI	population-based dataset	56	this study	59
13	cockroaches	Blaberidae	*Salganea* [1]	COII	22/50 species [C]	36	this study	33
14	net-winged beetles	Lycidae	*Scarelus* [1]	mtDNA	18/32 species	18	*	35
15	cockroaches	Blaberidae	several [1]	COII	21 species (9/10 genera)	22	this study	34
16	mite harvestmen	Stylocellidae	several [2]	mtDNA+nuDNA	98/300 putative species (36 species described)	95	this study / [1]	10
17	fig wasps	Sycophaginae	several [1]	mtDNA+nuDNA	55 spec (all genera)	55	#	12
18	prawns	Palaemonidae	*Macrobrachium* [3]	COI	population-based dataset	93	this study / [2]	13
19	prawns	Palaemonidae	*Macrobrachium* [3]	16S rRNA	45/105 species	43	this study	67
20	crabs	Gecarcinucidae	several [3]	mtDNA+nuDNA	61 species (55% genera, 60% species)	61	$	26
21	crabs	Potamidae	several [3]	16S	72 species (51% genera, 14% species)	65	this study (AAR)	57
22	freshwater bivalves	Corbiculidae	*Corbicula* [3]	COI	7/c. 30 species [D]	55	this study / [3]	64
23	freshwater snails	Pachychilidae	several [3]	16S rRNA	21/c. 140 species (all genera)	21	this study	28
24	freshwater snails	Pachychilidae	*Sulcospira* [3]	16S rRNA	57/c. 140 species (all genera)	129	this study / [3]	27
25	cyprinids	Cyprinidae	*Barbodes* [3]	CR	population-based dataset	19	this study	36
26	ricefishes	Adrianichthyidae	*Oryzias* [3]	mtDNA	13/24 species [E]	23	this study	60
27	cobitids	Cobitidae	*Pangio* [3]	Cyt b	18/32 species	77	this study	6
28	fanged frogs	Dicroglossidae	*Limnonectes* [1]	mtDNA	45/55 species [F]	78	this study	19
29	frogs	Ranidae	*Rana* [2]	ND3	14/? species (several cryptic species)	15	this study	58
30	frogs	Ranidae	*Rana* [1]	mtDNA	10/? species (several cryptic species)	54	this study	21
31	water snakes	Homalopsidae	several [3]	Cyt b	20/34 species	24	this study	1
32	spiderhunters	Nectariniidae	*Arachnothera* [1]	mtDNA	10/10 species	46	this study	41
33	bush warblers	Cettiidae	*Cettia* [1]	Cyt b	4/4 species (13/28 subspecies)	12	this study	48
34	kingfishers	Alcedinidae	*Ceyx* [1]	ND2	4/6 species [G]	19	this study	30
35	robins	Turdidae	*Copsychus* [1]	mtDNA	9/19 subspecies	33	this study	56
36	flowerpeckers	Dicaeidae	*Dicaeum* [1]	ND2	population-based dataset	16	this study	31

Table continued

Dataset[a]	Common name	Family	Genus [generalized habitat][b]	Genetic marker	Taxonomic coverage[f]	nTaxa (ingroup)[g]	BEAST/ Lagrange Performed[h]	Source Ref[i]
37	forktails	Muscicapidae	*Enicurus* [1]	mtDNA	population-based dataset	15	this study	40
38	flycatchers	Muscicapidae	*Ficedula* [1]	Cyt b	25/25 species	27	this study / [4]	49
39	cuckooshrikes	Campephagidae	*Lalage* [1]	ND2	population-based dataset	12	this study	31
40	spiderhunters	Nectariniidae	*Nectarinia* [1]	ND2	population-based dataset	30	this study	31
41	leaf-warblers	Phylloscopidae	*Phylloscopus* [2]	ND2	population-based dataset	19	this study	24
42	bulbuls	Pycnonotidae	*Pycnonotus* [1]	ND2	population-based dataset	46	this study	31
43	fantails	Rhipiduridae	*Rhipidura* [1]	ND2	population-based dataset	50	this study	31
44	megapodes	Megapodiidae	several [1]	ND2	15/22 species	23	this study	4
45	white-eyes	Zosteropidae	several [1]	mtDNA	42/c. 80 species ('clade B')	57	this study	39
46	whistlers	Pachycephalidae	several [1]	ND2	35/49 species	39	*	25
47	bulbuls	Pycnonotidae	several [1]	mtDNA	43/> 130 species [H]	46	this study	47
48	shrews	Soricidae	*Crocidura* [1]	Cyt b	27/27 species (Malay Archipelago) [I]	47	§	18
49	wild dogs	Canidae	*Cuon* [2]	CR	population-based dataset	19	this study	22
50	fruit bats	Pteropdidae	*Cynopterus* [1]	Cyt b	5/7 species [J]	140	this study	9 + 53
51	elephants	Elaphantidae	*Elaphas* [2]	mtDNA	population-based dataset	32	§	63
52	bats	Hipposideridae	*Hipposideros* [2]	ND2	21/70 species	57	this study	45
53	macaques	Cercopithecidae	*Macaca* [1]	mtDNA+nuDNA	5/20-22 species [K]	23	this study	62
54	macaques	Cercopithecidae	*Macaca* [1]	mtDNA	15/20-22 species [K]	17	this study	68
55	clouded leopards	Felidae	*Neofelis* [1]	mtDNA	2/2 species	7	this study	8
56	tigers	Felidae	*Panthera* [1]	mtDNA	population-based dataset	25	this study	32
57	palm civets	Viverridae	*Paradoxurus* [1]	mtDNA	population-based dataset	86	§	50
58	rodents	Muridae	several [2]	mtDNA+nuDNA	36 species (4/22 genera)	36	this study / [4]	23
59	squirrels	Sciuridae	several [2]	mtDNA	15 species (all genera)	15	this study / [4]	37
60	tree squirrels	Sciuridae	*Sundascirus* [1]	Cyt b	14/15 species	29	§	15
61	treeshrews	Tupaiidae	*Tupaia* [1]	mtDNA	20/20 species	18	§	54

[a] Datasets showing 'Out-of-Borneo' signatures are marked grey. [b] Generalized habitat type: 1 = rainforest, 2 = other habitat(s), 3 = freshwater. [c] Calibration (F) = fossil calibration, calibration (G) = geological calibration, calibration (F+G) = fossil and geological calibration, calibration (I) = indirect calibration points from other study. [d] Number of haplotypes/lineages/species according to source publication. [e] Root height inferred from BEAST analyses; numbers in square brackets denote root ages directly obtained from source publication. [f] Taxonomic coverage according to source authors; e.g., 71/113 species = 71 of 113 currently described species used in analysis; [g] Number of haplotypes (see source for information on lineages/species). [h] AAR/Lagrange information: * source publication uses Lagrange, DIVA, and MrBayes for AAR, respectively – data used for statistics and colonization routes in Fig. 1.4 and Fig. S14-15; # source publication performed AAR, but some defined areas summarize several islands (e.g., 'Sundaland'); § AAR not performed in source publication; $ AAR not performed in source publication but colonization routes inferred from topology; [1] not all terminal colonizations shown in Fig. 1.4 and Fig. S14 and S15, i.e. single lineages from Indochina, Thai-Malay Peninsula, Sumatra, and Borneo with uncertain ancestral area; [2] Lagrange not required, but topology suggests colonization from East of Wallacea to Sumatra in the Plio-Pleistocene (shown in Fig. 1.4 and Fig. S14), the remainder shows unresolved colonization routes between Indochina, Thai-Malay Peninsula, Sumatra, and Borneo; [3] Lagrange not feasible due to several unresolved basal relationships; [4] Lagrange not feasible due to several ambiguous distribution areas in certain taxa. [i] Source reference: see Supplementary Information for source references.

Table 1.2. List of datasets used for phylogenetic meta-analysis (plus ancestral area reconstructions) including information on taxa, substitution rates/fossil calibration applied by (source) authors, and results inferred from BEAST analyses (root height).

Dataset	Family	Genus	Substitution model used for analyses[a] part1/part2/part3	jModelTest selection AIC, BIC[a] part1/part2/part3	Rates (%/Ma)/ dates used for BEAST[b]	Root height (Ma): mean (95% HPD)[c]	Source ref[d]	Rates/dates source ref[d]
01	Meliaceae	*Aglaia*	G+G	G+G, G+G	calibration (F)	108.3 (70.8, 156.7)	42	see source
02	Araceae	*Alocasia*			calibration (F)	[47.5 (47.0, 48.0)]	46	see source
03	Begoniaceae	*Begonia*			calibration (F)	[22.3 (N.A., N.A.)]	61	see source
04	Gesneriaceae	*Cyrtandra*	H+G*	G+G | K80+G	0.1	167.9 (107.6, 245.2)	2	see source
05	Ericaceae	*Rhododendron*	G+G	G+G, G+G	calibration (F)	59.9 (58.0, 61.9)	20	38
06	Rafflesiaceae	several			calibration (I)	[96.1 (83.6, 110.7)]	3	see source
07	Culicidae	*Anopheles*	G+G/H+I+G	G+G/H+I+G, G+G/H+I	2.3	1.5 (0.9, 2.2)	16	7
08	Nymphalidae	*Cethosia*			calibration (F)	[67.0 (57.0, 82.0)]	43	see source
09	Nymphalidae	*Charaxes*			calibration (F)	[41.5 (38.5, 44.0)]	44	see source
10	Pieridae	*Delias*			calibration (I)	[N.A.]	69	see source
11	Lycidae	*Metriorrhynchus*			2.3	[N.A.]	5	see source
12	Tetragnathidae	*Nephila*	H+G	H+G, H+G	2.3	1.6 (0.9, 2.4)	59	7
13	Blaberidae	*Salganea*	$		3.8	12.6 (9.4, 16.1)	33	34
14	Lycidae	*Scarelus*			2.3	[N.A.]	35	see source
15	Blaberidae	several	H+I+G	H+I+G, H+I+G	3.8	17.2 (10.1, 24.9)	34	see source
16	Stylocellidae	several	$		calibration (F)	425.0 (423.0, 427.0)	10	see source
17	Sycophaginae	several			calibration (F+G)	[48.2]	12	see source
18	Palaemonidae	*Macrobrachium*	H+G	H+G, H+G	1.485	3.5 (2.2, 4.8)	13	66
19	Palaemonidae	*Macrobrachium*	H+G	H+G, H+G	0.745	44.3 (24.3, 65.9)	67	14
20	Gecarcinucidae	several			calibration (F)	54.5 (35.0, 76.6)	26	see source
21	Potamidae	several			calibration (F+G)	[N.A.]	57	see source
22	Corbiculidae	*Corbicula*	$		1.48	1.9 (1.1, 2.9)	64	66
23	Pachychilidae	several	$		1.0	78.7 (53.3, 106.8)	28	see source
24	Pachychilidae	*Sulcospira*	G+G[#]	H+G, H+G	1.0	13.2 (8.9, 18.0)	27	28
25	Cyprinidae	*Barbodes*	H+G	H+G, H+G	3.6	9.3 (2.3, 20.7)	36	see source
26	Adrianichthyidae	*Oryzias*	$		2.5	7.7 (5.3, 10.4)	60	17
27	Cobitidae	*Pangio*	G+I+G	G+I+G, H+I+G	0.68	48.0 (32.0, 66.4)	6	see source
28	Dicroglossidae	*Limnonectes*	$		1.48	22.8 (17.0, 29.5)	19	11
29	Ranidae	*Rana*	G+G	G+G, G+G	2.4	13.7 (9.3, 18.5)	58	52
30	Ranidae	*Rana*	H+I+G*	G+I+G, G+I+G	2.4	11.4 (7.7, 15.3)	21	52
31	Homalopsidae	several	$		calibration (F)	41.9 (29.6, 55.1)	1	see source
32	Nectariniidae	*Arachnothera*	H+I+G/H+I+G*	G+I+G/G+I+G, G+I+G/H+I+G	2.1	15.7 (10.8, 20.6)	41	65
33	Cettiidae	*Cettia*	H+G	H+G, H+G	2.1	11.5 (5.5, 18.0)	48	65
34	Alcedinidae	*Ceyx*	G+G	G+G, H+G	2.1	7.2 (5.0, 9.5)	30	65

Table continued

Dataset	Family	Genus	Substitution model used for analyses[a] part1/part2/part3	jModelTest selection AIC, BIC[a] part1/part2/part3	Rates (%/Ma)/ dates used for BEAST[b]	Root height (Ma): mean (95% HPD)[c]	Source ref[d]	Rates/dates source ref[d]
35	Turdidae	*Copsychus*	H+I+G/H+I*	G+I+G/G+I, H+G/G+I	2.1	1.8 (1.0, 2.7)	56	65
36	Dicaeidae	*Dicaeum*	G+G	G+G, G+G	2.1	7.9 (4.5, 11.4)	31	65
37	Muscicapidae	*Enicurus*	S+I+G/H+G	S+I+G/H+G, S+G/H+G	2.1	7.1 (4.4, 9.7)	40	65
38	Muscicapidae	*Ficedula*	G+G	G+G, H+G	2.1	7.7 (5.4, 10.6)	49	65
39	Campephagidae	*Lalage*	G+G	G+G, H+G	2.1	7.7 (2.9, 12.1)	31	65
40	Nectariniidae	*Nectarinia*	G+G	G+G, G+G	2.1	6.7 (4.9, 8.5)	31	65
41	Phylloscopidae	*Phylloscopus*	H+I+G	H+I+G, H+G	2.1	7.4 (3.8, 11.1)	24	65
42	Pycnonotidae	*Pycnonotus*	H+I+G*	G+I+G, H+G	2.1	8.4 (4.4, 12.8)	31	65
43	Rhipiduridae	*Rhipidura*	H+I+G*	G+I+G, G+I+G	2.1	9.6 (2.3, 24.8)	31	65
44	Megapodiidae	several	$		1.79	22.6 (14.8, 30.6)	4	51
45	Zosteropidae	several	G+I+G/H+I+G	G+I+G/H+I+G, G+I+G/H+I+G	2.1	22.7 (19.3, 26.1)	39	65
46	Pachycephalidae	several			2.1	[7.2 (4.1, 10.9)]	25	65
47	Pycnonotidae	several	G+I+G/H+G	G+I+G/H+G, G+I+G/H+G	2.1	20.1 (16.5, 23.7)	47	65
48	Soricidae	*Crocidura*			calibration (F)	[8.0 (7.2, 8.9)]	18	see source
49	Canidae	*Cuon*	H+G	H+G, H+G	5.48	1.4 (0.6, 2.5)	22	29
50	Pteropdidae	*Cynopterus*	H+I+G*	G+I+G, G+I+G	4.7	1.3 (0.8, 1.7)	9 + 53	55
51	Elaphantidae	*Elaphas*			calibration (F)	7.4 (7.3, 8.0)	63	see source
52	Hipposideridae	*Hipposideros*	G+G	G+G, G+G	4.7	8.2 (5.3, 11.1)	45	55
53	Cercopithecidae	*Macaca*	G+G/H+G/H	G+G/H+G/H, JC/H+G/K80	calibration (F)	7.3 (4.6, 10.3)	62	see source
54	Cercopithecidae	*Macaca*	H+G/H+G	H+G/H+G, H+G/H+G	calibration (F)	9.9 (7.1, 12.8)	68	see source
55	Felidae	*Neofelis*	H+G/H+G	H+G/H+G, H/H+G	calibration (F)	7.1 (5.4, 9.0)	8	see source
56	Felidae	*Panthera*	H+I+G*	G+I+G, H+G	1.53	1.6 (0.8, 2.8)	32	see source
57	Viverridae	*Paradoxurus*			calibration (F)	[N.A.]	50	see source
58	Muridae	several	G+I+G	G+I+G, G+I+G	calibration (F)	140.4 (99.5, 185.8)	23	see source
59	Sciuridae	several	$		calibration (F)	39.0 (34.5, 45.2)	37	see source
60	Sciuridae	*Sundascirus*			calibration (I)	[10.0]	15	see source
61	Tupaiidae	*Tupaia*			calibration (F)	[83.5 (68.0, 93.0)]	54	see source

[a] Substitution models: G = GTR, H = HKY, S = SYM; e.g., G+I = GTR+I, H+G = HKY+G, S+I+G = SYM+I+G; $ = dataset re-used from Stelbrink et al. (2012); * = less complex model used but also Gamma and Propinv parameters from AIC (jModelTest) due to low ESS values for parameters 'prior' and 'posterior' (see text for details); # model selected by AIC resulted in low basal branch supports and thus in a non-monophyly of the ingroup – therefore, GTR+G was used. [b] Calibration (F) = fossil calibration, calibration (G) = geological calibration, calibration (F+G) = fossil and geological calibration, calibration (I) = indirect calibration points from other study. [c] Root height inferred from BEAST analyses; numbers in square brackets denote root ages directly obtained from source publication. [d] Source reference: see Supplementary Information for source references.

Table 1.3. Number of colonizations found between two areas (redundant 'Out-of' colonizations in a single dataset are counted once only). Area codes in square brackets denote areas defined *a priori*: A = Indochina, B = Thai-Malay Peninsula, C = Sumatra, D = Philippines, PAL = Palawan, E = Borneo, F = Java, G = Wallacea excluding Sulawesi, SUL = Sulawesi, H = East of Wallacea. Top 20 are highlighted in gray. Note that total number of emigrations/ immigrations presented in Table S4 and tested statistically for significance are higher as they also include emigrations/immigrations to/from India/Sri Lanka and Japan (see Fig. 1.4).

Area-pair: source–destination	pre-Miocene	Miocene	Plio-Pleistocene	Total	Area-pair: source–destination	pre-Miocene	Miocene	Plio-Pleistocene	Total
A–B	1	4	10	15	E–A	1	9	2	12
A–C		3	7	10	E–B	4	6	11	21
A–D		4	6	10	E–C	1	3	8	12
A–PAL			1	1	E–D	3	6	5	14
A–E		6	8	14	E–PAL	1	3	2	6
A–F		2	4	6	E–F		5	4	9
A–G		1	1	2	E–G		2		2
A–SUL		3	6	9	E–SUL	1	3	2	6
A–H		3	3	6	E–H	1	4		5
B–A	1	1	3	5	F–A		1	1	2
B–C	2	1	3	6	F–B			1	1
B–D	1			1	F–C	1	1		2
B–PAL				0	F–D				0
B–E	1	1	3	5	F–PAL				0
B–F				0	F–E				0
B–G		1	1	2	F–G		1		1
B–SUL	1	1		2	F–SUL				0
B–H				0	F–H	1			1
C–A		1	1	2	G–A		1		1
C–B	1	2	1	4	G–B				0
C–D			2	2	G–C			1	1
C–PAL				0	G–D			2	2
C–E		1	5	6	G–PAL				0
C–F	1	1	5	7	G–E			1	1
C–G				0	G–F				0
C–SUL		1		1	G–SUL		1		1
C–H				0	G–H			4	4
D–A			3	3	SUL–A		1		1
D–B			1	1	SUL–B		1		1
D–C		1		1	SUL–C				0
D–PAL		2	3	5	SUL–D		1	1	2
D–E		1	2	3	SUL–PAL		1		1
D–F			1	1	SUL–E				0
D–G		1		1	SUL–F		1		1
D–SUL		3	3	6	SUL–G				0
D–H		2	4	6	SUL–H	1	1	2	4
PAL–A				0	H–A			1	1
PAL–B				0	H–B				0
PAL–C				0	H–C		1	1	2
PAL–D	1			1	H–D			5	5
PAL–E				0	H–PAL				0
PAL–F				0	H–E				0
PAL–G				0	H–F		1		1
PAL–SUL		1		1	H–G		1	2	3
PAL–H	1			1	H–SUL	1	3	2	6

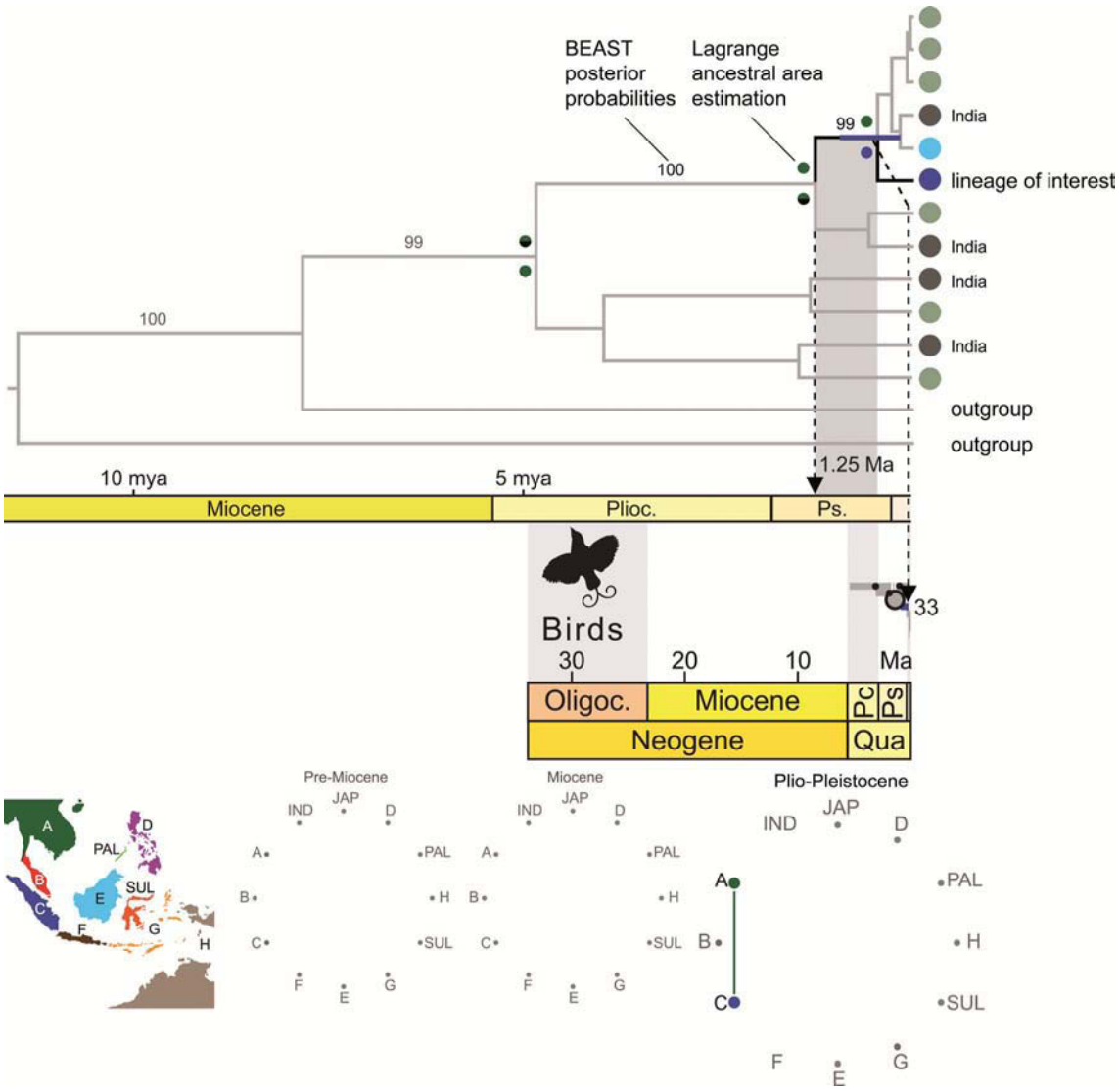


Figure 1.3. Exemplary illustration (dataset 33, bush warblers) how colonization routes through time were inferred from both MCC trees (BEAST) and ancestral area estimation (Lagrange; colored circles). The blue time bar (0.16-0.9 Mya) represents the divergence time of the MRCA in Sumatra and is plotted in Fig. S11 and S13 (middle panel). The lineage of interest represents a specimen from area C (Sumatra). This lineage can be traced back to a node with an ancestral area in A (Indochina) 1.25 Mya (black arrow) and hence a colonization route from area A to area C in the 'Plio-Pleistocene' timeframe is visualized in the bottom panel (see Fig. 1.4 and Fig. S14 and S15).

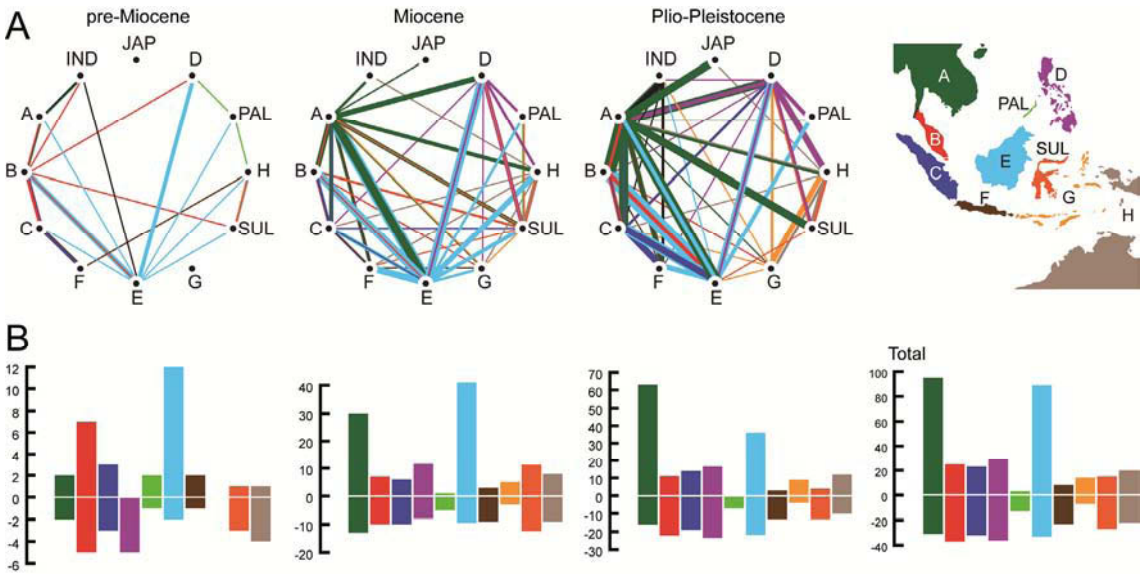


Figure 1.4. Colonization routes of extant Southeast Asian taxa. A) Colonization routes inferred from ancestral area estimation (DEC model, Lagrange) for lineages among studied taxa in the pre-Miocene, Miocene, and Plio-Pleistocene (see Materials and Methods and Fig. 1.3 for details). Line-color corresponds to inferred area of origin (see inset map), while line thickness corresponds to number of colonizations; multiple colonizations found for a specific source and destination (e.g., Indochina-Sumatra) from a single dataset are counted once only. See Fig. S14 and S15 for group-specific colonization routes. B) Number of emigration (positive bars) and immigration (negative bars) events observed for pre-defined areas in the pre-Miocene, Miocene, and Plio-Pleistocene.

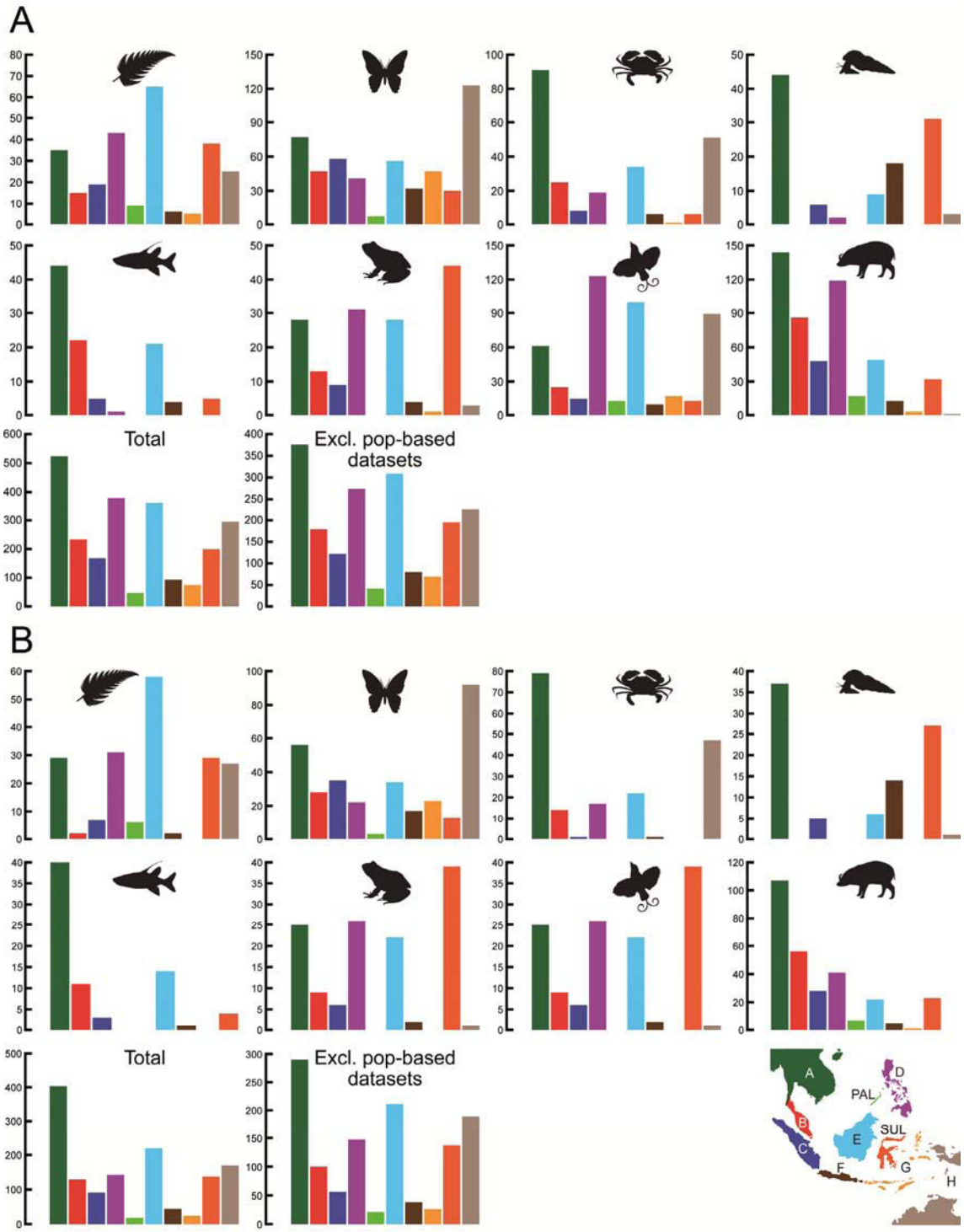


Figure 1.5. Distribution of lineages and *in situ* diversification events among all 61 datasets through time. A) Group-specific (plants, insects & spiders, freshwater crustaceans, freshwater molluscs, freshwater fishes, herpetofauna, birds, and mammals), total combined, and total combined excluding population-based number of lineages across all datasets from each pre-defined area. B) Group-specific (as above), total combined, and total combined excluding population-based number of *in situ* diversification events across all datasets from each pre-defined area.

Mammal and plant species richness across SE Asia

Comprehensive distribution data exists for SE Asian plants and mammals, allowing comparison of extant species richness of these groups across our pre-defined areas to that recovered from our phylogenetic and ancestral area analyses. To construct a map of mammal species richness, we collected information on the current distribution of all mammal species occurring in the study area through the Global Mammal Assessment (Rondinini et al., 2011). Considering an area ranging from Myanmar to New Guinea, we covered a total of 1,086 mammal species, belonging to 17 orders (Catullo et al., 2008); Rodentia (number of species = 388), Chiroptera (328), Primates (76), Diprotodontia (58), Soricomorpha (56), Carnivora (54), Artiodactyla (52), Scandentia (17), Dasyuromorphia (16), Peramelemorphia (12), Erinaceomorpha (7), Lagomorpha (6), Perissodactyla (3), Pholidota (3), Dermoptera (2), and Proboscidea (1). For each species we obtained the most updated available global distribution range. We weighted each species according to the inverse of its area of distribution (i.e., species presence is divided by the square kilometres of their distribution ranges) and calculated a map of species richness (10 arcminutes spatial resolution) (Fig. 1.6). Endemic species and species mostly abundant in Southeast Asia, 'locally restricted taxa', with a high number of specimens in a comparatively small area will result in a higher weighted value, and are thus considered more important in the analyses.

Data for vascular plants was used from Kreft & Jetz (2007) to create a map (Fig. 1.6) of species richness, with permission (*Copyright (2012) National Academy of Sciences, U.S.A.*). Considering 1,032 geographic units worldwide and a set of bioclimatic and environmental variables, the source authors generated a set of predictions for species richness (1 degree resolution) (Kreft & Jetz, 2007). Among all models of species richness available from Kreft & Jetz (2007), we considered the kriging estimate here, as it was demonstrated that the inclusion of neighbourhood effects substantially improve the quality of predictions.

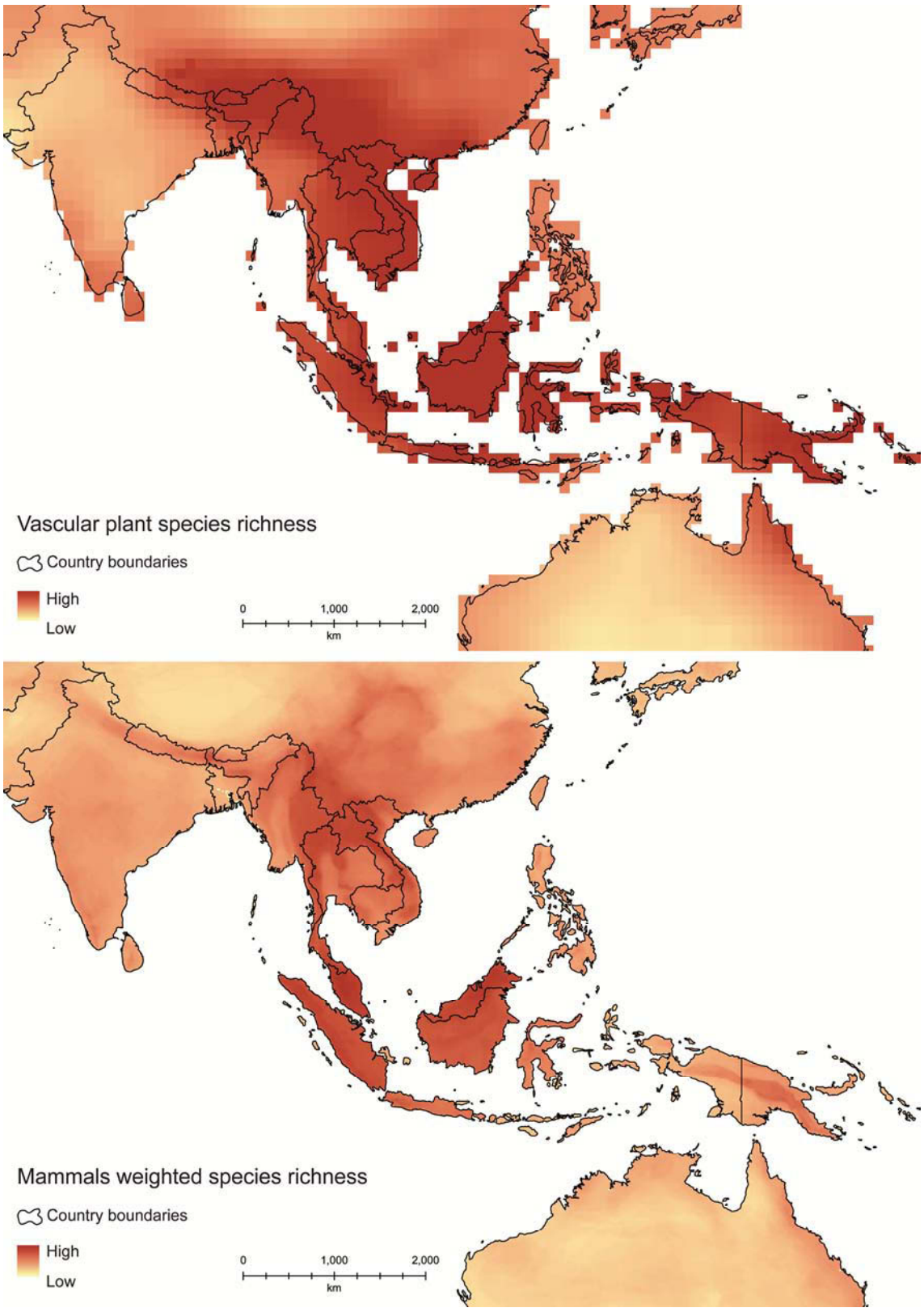


Figure 1.6. Extant Southeast Asian mammal and plant species richness calculated as the sum of species-specific distribution models. Borneo hosts the highest extant number of mammalian species (n = 230), and also the highest weighted species richness for both vascular plants and mammals (together with Thai-Malay Peninsula for top 10% of mammals) for SE Asia (see Materials and Methods for details). Plant species richness data from Kreft & Jetz (2007), *Copyright (2012) National Academy of Sciences, U.S.A*.

Results

Setting the Scene: Palynological and geological syntheses suggest major wet refugia in Indochina and Borneo

The early Miocene collision of Australia with the eastern margin of Sundaland dramatically changed the region's palaeogeography (Hall, 2009b). This collision resulted in the formation of widespread uplands in central Borneo from the early Miocene onward, and the development of major rivers, such as the Mahakam and Baram, and their respective deltas. The Meratus Mountains (SE Borneo) were uplifted during the middle and late Miocene (Fig. 1.1, 15 and 10 Mya) (Witts et al., 2012) and Mount Kinabalu during the late Miocene (Hall, 2011). Widespread evergreen rainforests would have covered much of Sundaland during the early and middle Miocene. There was some climate variability, with drier climates being widespread during periods of low sea level at subequatorial latitudes, and with rain forests reaching their maximum northward extent during the middle Miocene. However, whereas palynological data from the late Miocene and Pliocene suggest periodicity of climate in the region of the Sunda Shelf to the west, high resolution palynological data (Supplementary Information) back to at least 9 Mya from East Borneo suggest continued everwet climates, characterizing periods of both high and low (glacial) sea level in the equatorial region (Morley & Morley, 2011). This climatic stability has been invoked as an explanation for high floristic diversity witnessed in areas such as Lambir Hills in Sarawak (Borneo), which harbours rain forests rivalling those of the Neotropics. The northern Sunda Shelf was subject to extensive marine inundation, with the development of widespread inland shallow seas, reaching their greatest extent by the middle Miocene thermal (and sea-level) maximum (Fig. 1.1, 15 Mya).

Phylogenetic meta-analyses: Origins of SE Asian biota

Our meta-analyses of published molecular datasets using a standardized relaxed Bayesian dating approach (Supplementary Information) identified a history of phylogenetic lineage diversification in SE Asian biota since the Jurassic (Fig. S5-S13, Table S1 for detailed results). Only two of our study groups date back to the Mesozoic (plants, insects & spiders). While SE Asia's continental core – Indochina and Sundaland – was essentially established by the end of the Mesozoic, its current largely insular topography and megadiverse biota were decisively shaped by Cenozoic events. For the vast majority of phylogenetic datasets examined, diversification events were Cenozoic in age (Fig. S5-S13, Table S1 for detailed results). Geographic and biotic changes are examined here in light of the interplay of tectonics and climatic oscillations, most notably evident during three distinctive stages in the Paleogene (~65.5-23 Mya), Neogene (~23-2.6 Mya), and the Quaternary (~2.6 Mya-present; see timeline in Fig. S13).

Borneo and Indochina – major 'evolutionary hotspots'

Diversity can be evaluated here as the number and variety of lineages found in a specific pre-defined area (Table 1.1). This diversity may have arisen through the accumulation of immigrants, by *in situ* diversification, or some combination of the two. Indochina, the Philippines and Borneo held the highest total number of lineages across all taxonomic groups (n>300; Fig. 1.5A), which were statistically significantly higher than the other pre-defined areas: 'East of Wallacea' (n = 295), the Thai-Malay Peninsula (n = 233), Sulawesi (n = 199), Sumatra (n = 168), Java (n = 93), 'Wallacea excl. Sulawesi' (n = 74), and Palawan (n = 46) (Mann-Whitney U test: $p<0.05$; Table S4); other significant differences in pairwise comparisons are due to low numbers. These differences remained when population-based datasets were excluded (Fig. 1.5A, Table S4).

Significantly higher numbers of *in situ* diversification events were observed in several areas: Indochina compared with all areas except 'East of Wallacea' (see Fig. 1.5B; $p<0.05$; Table S4); Borneo compared with Sumatra, Palawan, Java and 'Wallacea excl. Sulawesi' ($p<0.05$; however, this can primarily be attributed to particular taxonomic groups, namely plants, fishes, herpetofauna, and birds); and Sulawesi compared with Palawan and 'Wallacea excl. Sulawesi' for most taxon groups ($p<0.05$). However, inclusion of population-based datasets (see Table 1.1) increased both the number of total lineages and diversification events. The removal of population-based datasets reduced significant differences to some extent, particularly for Indochina (Fig. 1.5; Table S4).

Indices for total number of lineages and *in situ* diversification events for each pre-defined area (excl. population-based datasets) were calculated in an attempt to control for potential biases resulting from increased lineage sampling for a particular area. These indices were tested for significant differences between areas (Mann-Whitney U tests, Table S4). A significantly higher proportion of *in situ* diversification events were still observed for Indochina (*vs*. Thai-Malay Peninsula, Sumatra, Palawan, Java, and 'Wallacea excl. Sulawesi'; $p<0.05$) and Borneo (*vs*. Java; $p<0.05$), when considering total number of lineages sampled per area (excl. population-based datasets).

Because statistically significant higher total lineage numbers and number of *in situ* diversification events for Indochina and Borneo could simply have been a function of area size, an area size-dependent index for all taxonomic groups among all areas was calculated for these parameters (Table 1.4). Interestingly, the comparatively large current size of both Indochina and Borneo was not reflected by higher indices reflecting size-dependent diversity (total number of lineages and number of *in situ* diversification events). A comparatively lower diversity-size index was found for the biogeographic area 'East of Wallacea', likely related to under-sampling of Sahul Shelf representatives. In contrast, Sulawesi and the Philippines – both biogeographic areas with comparatively high numbers of lineages and *in situ* diversification events – showed

Table 1.4. Total numbers of lineages, *in situ* diversification events, and number of mammal species related to area size. Area codes denote biogeographic areas defined *a priori*: A = Indochina, B = Thai-Malay Peninsula, C = Sumatra, D = Philippines, PAL = Palawan, E = Borneo, F = Java, G = Wallacea excluding Sulawesi, SUL = Sulawesi, H = East of Wallacea.

Biogeographic area	Area [km^2][a]	N lineages[b]	Lineages / 1,000 km^2	N *in situ* diversifications[b]	Diversifications / 1,000 km^2	N Mammals species[c]	Mammals species / 1,000 km^2	N emigrations	Emigration / 1,000 km^2
A	1,938,743	376	0.19	290	0.15	453	0.23	95	0.05
B	131,598	179	1.36	100	0.76	67	0.51	25	0.19
C	473,481	122	0.26	56	0.12	97	0.20	23	0.05
D	285,350	274	0.96	147	0.52	150	0.53	29	0.10
PAL	14,650	41	2.80	21	1.43	58*	3.96	3	0.20
E	743,330	309	0.42	211	0.28	190	0.26	89	0.12
F	138,794	80	0.58	38	0.27	59	0.43	8	0.06
G	142,090	69	0.49	26	0.18	22	0.15	14	0.10
SUL	174,600	195	1.12	137	0.78	51	0.29	15	0.09
H	6,072,464	227	0.04	189	0.03	101	0.02	20	<0.01

[a] Present-day area as in Fig. 1.6. [b] Total numbers of lineages and *in situ* diversification events excluding population-based datasets. [c] Source: *Esselstyn J.A., Widmann P., Heaney L.R. 2004. The mammals of Palawan Island, Philippines. Proc. Biol. Soc. Washingt. 117:271–302; otherwise: Global Mammal Assessment program: Rondinini C., Di Marco M., Chiozza F., Santulli G., Baisero D., Visconti P., Hoffmann M., Schipper J., Stuart S.N., Tognelli M.F., Amori G., Falcucci A., Maiorano L., Boitani L. 2011. Global habitat suitability models of terrestrial mammals. Philos. Trans. R. Soc. London B 366:2633–2641.

comparatively high indices for their relatively small geographic sizes. Similarly, Palawan, the smallest area examined here showed the highest indices for these two parameters (Table 1.4). One important caveat on these findings is that current area size is not necessarily representative of the extended timeframe under investigation here.

Indochina and Borneo have been major colonization sources reflected by a significantly higher number of emigration events through all time periods examined here compared with our other pre-defined areas (see Fig. 1.4; Mann-Whitney *U* test: $p<0.01$ for Indochina and $p<0.05$ for Borneo; Table S4). Thirty-two of 49 datasets analyzed for colonization routes show 'Out-of-Borneo' signatures (total number of emigrations: $n = 89$; total number of emigrations excl. India/Sri Lanka and Japan: $n=87$; compare with Fig. 1.4 and Fig. S14 and S15), while 36 datasets show colonizations out of Indochina (total number of emigrations: $n = 95$; total number of emigrations excl. India/Sri Lanka and Japan: $n = 73$; Table 1.3 and Table S3). An increased total number of emigrations were observed in the Miocene, and even more pronounced in the Plio-Pleistocene (pre-Miocene: $n = 30$, Miocene: $n = 122$, Plio-Pleistocene: $n = 169$), the latter significantly different to that of the pre-Miocene (Mann-Whitney *U* test: $p<0.01$; Table S4). In contrast, immigration events are comparably much less frequent for these two areas (Indochina: $p<0.01$, Borneo: $p<0.05$), with other areas showing a relatively balanced or even negative emigration-immigration index, e.g. Palawan and Java (Fig. 1.4, lower panel; not significant except for Java $p<0.05$; Table S4). Interestingly, and as expected, very few emigration events were observed from younger volcanically active emergent areas such as Java ($n = 7$) (Fig. 1.4, Fig. S14 and S15). Despite their geographical proximity, colonizations of Sulawesi from Borneo have been infrequent ($n = 6$; Fig. 1.4, Fig. S14 and S15). The total number of immigration events found across all areas is significantly different between the three timeframes, increasing towards the present (pre-Miocene: $n = 26$; Miocene: $n = 86$; Plio-Pleistocene: $n = 149$; pre-Miocene *vs.* Miocene: $p<0.01$; pre-Miocene *vs.* Plio-Pleistocene: $p<0.001$; Miocene *vs.* Plio-Pleistocene: $p<0.05$; Table S4).

A pairwise comparison of the number of datasets observed for a particular source-destination route again highlighted Borneo and Indochina as predominantly 'sources' rather than 'destinations' across SE Asia through time, e.g.: Borneo–Thai-Malay Peninsula: 21↔5, Borneo–Philippines: 14↔3, and Indochina–Thai-Malay Peninsula: 15↔5 (source↔destination number of datasets; Table 1.3). Further, the rate of colonizations between currently adjacent islands/landmasses is generally higher compared to between currently remote areas (Table 1.3; see also Fig. 1.4, Fig. S14 and S15). However, even 'East of Wallacea' appears to have been colonized to a considerable extent from Indochina and Borneo, even though geographic distances separating these areas have remained comparatively large throughout the timeframe under investigation (Fig. 1.4). The majority of colonization events were assigned to the Plio-Pleistocene, and most frequently between areas formerly adjoined via 'Sundaland' (Fig. 1.4,

Table 1.3). To control for total number of lineages sampled per area (sampling effect), we calculated an index for emigrations and total lineages (excl. population-based datasets) (Table S4). These indices were tested for significance between areas (Mann-Whitney *U* tests). Interestingly, significant differences were only found between Sulawesi *vs.* Indochina, Sumatra, and Borneo (lower indices for Sulawesi); and Borneo *vs.* Java (lower indices for Java; Table S4). This result indicates that the significantly higher levels of emigration events identified for Indochina and Borneo are clearly related to the higher diversity (represented by number of lineages) found in these areas. Elevated levels of emigration events found in Indochina and Borneo might actually relate to the size of the respective source area. However, size-dependent indices were not comparatively elevated for either of these biogeographic areas (Table S4), although current area size was used for these analyses and area size has changed considerably through time for several areas (Fig. 1.1).

Performing multiple pairwise statistical tests can include amendments such as the (sequential) Bonferroni correction, i.e., effectively dividing the p-value cutoff by the number of tests performed or by identifying rank-based Bonferroni-corrected p-value cutoffs. For the majority of pairwise comparisons made (e.g., number of lineages, in situ diversifications, emigrations) the number of tests was $n = 45$. Applying a Bonferroni correction resulted in a decreased p-value cutoff of 0.0011 and therefore a greatly reduced number of significant differences observed (see Table S4). However, several drawbacks have been identified for such corrections (see e.g., Perneger, 1998; Moran, 2003; Armstrong, 2014), thus, here we base our interpretations on non-corrected significance testing – effectively prioritising each individual area-pair test (see e.g., Armstrong, 2014).

Mammal and plant distributional datasets: Elevated diversity on Borneo

A meta-analysis of mammalian and plant distributional datasets, for which high-resolution data across SE Asia was available (Fig. 1.6), again identified Borneo, in particular, as an area of special biodiversity significance for SE Asia (Fig. 1.6). For our pre-defined areas, Borneo hosts the highest number of weighted (see Materials and Methods) terrestrial mammalian ($n = 230$) and vascular plant species richness in SE Asia, while the top 10% of weighted mammalian species richness is hosted in the northeast of the island and in the Thai-Malay Peninsula ($n = 228$; Fig. 1.6). This pattern was consistent for Borneo for total number of plant lineages identified here (Fig. 1.5A and Table S4), but was not the case for mammals. For mammalian lineages, Borneo ranked fourth after Indochina, the Philippines and the Thai-Malay Peninsula. Interestingly, the Philippines showed a marked contrast between present-day mammal species richness (comparatively low) and total number of mammalian lineages (comparatively high) (cf. Fig. 1.5A and Fig. 1.6). Size-dependent indices of general diversity across all taxonomic groups were additionally applied to present-day mammal species richness (Fig. 1.6). However, when

current island size is taken into account, mammal diversity (total number of species) is comparatively low for Indochina and Borneo, while it is comparatively high for Palawan, the Thai-Malay Peninsula, the Philippines and Java (Table S4).

Palynological and palaeontological syntheses: Middle to late Quaternary Sunda Shelf sea-level dynamics

There are three predominant sea-level regimes for the middle to late Quaternary Sunda Shelf (Fig. 1.2) (Morley, 2012). The most commonly recurring scenario (±55% of last million years, Fig. 1.2B) is of periods with sea levels 40-50 m below current levels, around half the current shelf emergent, and evergreen rainforests extending from Borneo to Sumatra. The second most common scenario (±37% of last million years, Fig. 1.2A) is of periods with very low sea levels, such as the LGM. Seasonal climate vegetation was very widespread, though may not have formed a continuous north-to-south corridor for every glacial maximum, and the exposed sandy soils of the shelf may have acted as a substantial barrier to dispersal (Slik et al., 2011). Seasonality of climate may have varied between glacial maxima, suggested by the occurrence of certain mammalian fossils in Java. Faunas from the penultimate and older glacials included many large mammals requiring open woodland (van den Bergh et al., 2001), whereas there were no such immigrants during the LGM for non-forest species. This suggests that opportunities for migration across the Sunda Shelf may have been inhibited during the LGM, while more open vegetation types, perhaps with a corridor of semi-evergreen forests, but without true savanna, may have been present during earlier glacials (Cannon et al., 2009). Examination of floristic elements in Java suggests that there are semi-evergreen elements that are common to Java and Indochina, but there are no clear Indochinese deciduous forest elements in Java. This suggests that there may have been a dispersal pathway for semi-evergreen elements during the LGM but not for elements requiring more seasonality of climate. From the faunal perspective, the mammals from the Javanese last glacial do not include big non-forest browsers except elephants. Again, this suggests limited dispersal opportunities for open-vegetation browsers. However, the presence of large open-vegetation browsers from earlier glacial intervals in Java suggests that more opportunities for browsers may have occurred earlier in the Quaternary, but due to the floristic differences with Indochina it is thought unlikely that there was a continuous 'savanna corridor'. The third climatic scenario is represented by the present day (Fig. 1.2C), with high sea levels and evergreen rainforests extending from the Kra Isthmus to West Java, including Borneo, but importantly, occurring for just 8% of the last million years, emphasizing the 'refugial' nature of present day SE Asian rainforests and fauna (Cannon et al., 2009).

These glacial events are thought to have facilitated repeated biotic range expansions between Sumatra, the Thai-Malay Peninsula, Borneo and Java during low sea levels, followed by vicariance as sea levels subsequently rose (Cannon et al., 2009; Gower et al., 2012; de Bruyn

et al., 2013a). While fossil data could contribute to our understanding of these dynamics, the terrestrial palaeontological record from Quaternary SE Asia is still very poorly understood. Though Quaternary sites are relatively abundant throughout the region (Louys & Meijaard, 2010), these are mostly represented by cave sites, with often poorly constrained ages. In addition, many taxa have not yet been studied by specialists, and much controversy remains regarding taxonomic delineation. Borneo and Sumatra have yielded few terrestrial fossil remains apart from the famous late Pleistocene Niah and Madai Cave sites. There have been a few isolated Proboscidean finds from Samarinda, Brunei and Sarawak, but the exact stratigraphic origin is invariably poorly unknown (Hooijer, 1952; Cranbrook, 2010). On the other side of 'Wallace's Line', in the realm of Wallacea, increasing efforts in fossil vertebrate studies were carried out during the second half of the 20th century, culminating in the discovery of an endemic late Pleistocene hominin, *Homo floresiensis*, on the oceanic island of Flores (Brown et al., 2004).

The majority of datasets examined here show 'Out-of-Borneo' and 'Out-of-Indochina' signatures for the Plio-Pleistocene period, while colonization routes from both Sumatra and the Thai-Malay Peninsula were substantially fewer (Fig. 1.4, Table 1.3). Borneo and Indochina have thus been *the* major refugia for Sundaland over at least the last 5 Ma, and likely earlier (see Neogene discussion above).

Discussion

Given the taxonomic coverage of SE Asia, our phylogenetic meta-analyses of diverse faunal and floral groups identified initiation of diversification during the Jurassic in plants, and insects & spiders, while all other groups examined diversified later during the Cenozoic (Fig. S5-S13, Table S1), consistent with the fossil record. Fossils indicate that a diverse mammalian fauna existed during the Eocene and Oligocene, including various primates, flying lemurs, mustelids, ruminants, and rodents in Myanmar and Thailand (e.g., Marivaux et al., 2004, 2006; Peigné et al., 2006; Beard et al., 2009). There are, however, few Paleogene mammal fossils from island SE Asia, and some of these are of uncertain provenance (e.g., an Eocene anthracothere found on Timor Island and a putative member of the artiodactyl family Haplobunodontidae found in western Borneo; see Ducrocq, 1996).

Interestingly, a few datasets show pre-Oligocene diversification events from the Philippines (excluding Palawan) and Java. However, given the small number of taxa with such patterns, dating issues, extinction and range evolution seem more likely at present than a major conflict with the geological evidence. Colonization routes and ancestral areas inferred from meta-analyses of phylogenetic data show that Indochina and Borneo were already major evolutionary hotspots and sources of diversification in the pre-Miocene (Fig. 1.4, Fig. S14 and

S15, Table 1.2 and 1.4). These results most likely reflect the palaeogeographic history of Borneo and Indochina. Borneo is the largest landmass of the former Sundaland, and was less affected by sea-level changes compared to other islands like Sumatra and Java, which were largely submerged during the Miocene or even later (Fig. 1.1). The presence of extended emergent areas including extensive rainforest (Fig. 1.2) through all time-periods examined here may explain the elevated levels of *in situ* diversification and emigration observed for Indochina and Borneo (Fig. 1.5, see below).

Caveat on phylogenetic meta-analyses

Only a few colonizations from the islands within 'Wallacea excl. Sulawesi', Palawan and the area east of 'Wallace's Line' (i.e., mainly New Guinea and Australia) were inferred conducting ancestral area reconstructions (Fig. 1.4, Fig. S14 and S15), consistent with a recently published meta-analysis focusing on Sulawesi and Wallacea, which identified that the majority of colonization events were derived from Sundaland, not within or east of Wallacea (Stelbrink et al., 2012). It seems very unlikely that comparatively small and young islands such as the Lesser Sunda Islands and Palawan might have played a major role as refugia and/or colonization sources throughout the timeframe under investigation here, when comparing them with larger, older landmasses, such as Borneo. However, one might assume that the lack of particular dispersal routes is a result of weak taxonomic coverage, because some islands are more difficult to access and/or research permits are difficult to procure and hence taxa from these areas are missing in several phylogenetic studies. One area that is likely to be under-represented in phylogenetic studies is New Guinea, and current evidence suggests this landmass has played a significant role in diversification dynamics for the wider region (e.g., Balke et al., 2009; Jønsson et al., 2011; Müller et al., 2013). Similarly, two recent herpetological studies suggest that Palawan may have been more important in the biogeographic history of SE Asia than generally recognized, acting as a raft for mainland SE Asian lineages when separating from the mainland c. 30 Mya, and thus enabled dispersal across the Philippines, Borneo, Sulawesi and the Thai-Malay Peninsula during the Miocene (so-called 'Palawan Ark Hypothesis'; Blackburn et al., 2010; Siler et al., 2012). However, for most Philippine species groups, Palawan apparently did not function as a dispersal gateway between Borneo and the rest of the Philippines (see e.g., Brown et al., 2013; cf. also Fig. 1.4). Some species likely dispersed from Borneo to the Philippines (except Palawan) through the Sulu Archipelago (Oliveros & Moyle, 2010), while others dispersed from Indochina, probably via Taiwan (Esselstyn & Oliveros, 2010). Finally, 15 of 61 datasets analyzed here are based on intraspecific data, while the remainder focus above the species level (Table 1.1). This could result in relatively young divergence times for these intraspecific datasets, compared with those datasets studying taxa of a higher taxonomic level (interspecific, intergeneric).

Caveat on plant and mammal distributional modelling

SE Asia is a global hotspot for mammal species richness (Catullo et al., 2008; Rondinini et al., 2011). Roughly one quarter of global mammal taxa occur in this area, with many new families and species, which have only been discovered recently (Jenkins et al., 2005; Musser et al., 2005). The region is also a global centre of plant richness, especially in northern Borneo and New Guinea. Overall, our study area includes 4 biodiversity hotspots (Sundaland, Wallacea, Philippines, Indo-Burma) as defined by Myers et al. (2000), more than 56,000 plant species are present in the area, of which 52% are endemic to the 4 biodiversity hotspots mentioned (Myers et al., 2000).

We caution against an over-interpretation of the species richness map (Fig. 1.6), as both are probably overestimating actual species distribution. In fact, increasingly, lowland forest areas in SE Asia are characterized by the so-called 'empty forest syndrome', with large animals (mainly primates, carnivores and ungulates) that are extinct in vast areas of their former distribution range because of commercial hunting, even if suitable habitat is still present (Redford, 1992; Milner-Gulland et al., 2003; Corlett, 2007). Similar issues characterize also the model of plant species richness (Fig. 1.6). SE Asia is characterized by the highest rate of deforestation globally (Achard et al., 2002; Wilcove et al., 2013), and primary vegetation losses equate to 92.2% of the original extent for Sundaland, 85% for Wallacea, 97% for the Philippines, and 95.1% for Indo-Burma (Myers et al., 2000) in recent decades. If all forest types are included, some 50% of Borneo remains forested (Miettinen et al., 2011) but annual losses continue (1.3%/year, Miettinen et al., 2011; 3%/year for peat swamp forests only, Wilcove et al., 2013). Given that the plant species richness map has been calculated from a database using statistical models, data points, and bioclimatic variables without considering recent annual deforestation and other variables that are likely important determinants of species distributions (e.g., species interactions, microclimatic conditions, etc.), the output map may overestimate the current vascular plants species richness. Importantly though, if absolute values were too high, the proportion and thus species diversity rankings for each area should remain consistent.

Moreover, the two maps of species richness (Fig. 1.6) actually provide different information. While the model of plant species richness gives a 'biogeographic' estimate of the potential species richness, mammal species richness is more similar to the actual species richness due to issues related to the historical alteration of the species distribution ranges (e.g., habitat destruction, hunting, and other human related activities may have altered the original species distribution). Thus, species richness for mammals may also reflect historical/human-related factors, as well as biogeography.

Indochina and Borneo: Major evolutionary hotspots

Geological and climatic processes associated with the early Miocene collision (see above) may have effectively divided the biota into two major Miocene areas north and south of the Thai-Malay Peninsula: namely, Indochina and (an extended) Borneo (Fig. 1.1, 15 Ma). Consistent with an expectation of *in situ* diversification in refugia and thus the accumulation of lineages during the global middle Miocene Climate and sea-level maximum, our analyses of phylogenetic diversification events show an increase in the rate of diversification consistent with this time period (Fig. 1.5 and Fig. S13, Table S4). This is most evident for Indochina and Borneo, suggesting a role for global Miocene sea-level fluctuations and associated refugia development on biotic evolution.

SE Asia's biodiversity has evidently been shaped by both geological and climatic drivers resulting in an extremely complicated history of past land (re)connections, including micro-terrane movements, with a young age of many diversification events evident in the phylogenetic meta-analyses, a high proportion of which occur in the latest Pliocene and Pleistocene (Fig. S13; see also Stelbrink et al., 2012 for details on Sulawesi). During the Pliocene, Borneo essentially maintained its earlier configuration. The Barisan Mountains of Sumatra continued to develop to their present altitudes (Fig. 1.1, 5 Ma); the Sumatran coastal plain expanded and at times of low sea level may have been contiguous with the Thai-Malay Peninsula. The island of Java essentially formed during this period and floristic differences from Sumatra as emphasized by van Welzen et al. (2011) probably relate to its persistently drier climate and young geological age. A substantial increase in the colonization of other Greater Sunda Islands and the Philippines from Borneo and Indochina in the Miocene and even more pronounced in the Plio-Pleistocene (Fig. 1.4 and Table 1.3), although representing a much shorter timeframe compared to the Miocene. An increase in colonization events between adjacent areas of Sundaland (see Fig. 1.4 and Table 1.3), is potentially contemporaneous with the successive (re)emergence of land in these areas and cyclical climate-induced connectivity (Fig. 1.2). Consistent with these young emergent ages, ancestral area estimation generally show less frequent colonization events out of Java and Sumatra (compared to Borneo and Indochina), but higher levels of immigration events into these areas (Fig. 1.4 and Table 1.3). Hence, climate-forced sea-level changes facilitating inter-area connection-disconnection events have played a substantial role in providing new dispersal 'corridors' for such lineage dispersal across SE Asia.

Indochina, and Borneo especially, the two largest SE Asian areas examined here with the longest emergent histories, have been major diversification hotspots through time and key sources for lineage dispersal across the region (Fig. 1.4, Table 1.3). Such larger areas may also have experienced comparatively fewer extinction events due to presence of additional putative refugia. Our analyses suggest length of emergent history has been a more important factor than

current area size for both total number of lineages and levels of emigration events for Indochina and Borneo (Table S4), however, future analytical approaches should consider changes in area size since emergence. For example, analytical advances to test for both ancestral area estimation and geography-diversification correlations have recently become available (e.g., R package BioGeoBEARS, Matzke, 2013; GeoSSE, Goldberg et al., 2011). Moreover, our analyses show that the prolonged emergent history of Indochina and Borneo initially facilitated the accumulation of lineages in both biogeographic areas (Fig. 1.5 and Fig. S13, Table S1 and S5), resulting in increased levels of emigration across SE Asia when remaining islands (re)emerged over the course of Mio- and Plio-Pleistocene sea-level fluctuations (Fig. 1.2 and 1.4, Table 1.3 and Table S3).

In summary, we find evidence for all four hypotheses set out in the Introduction, namely: levels of *in situ* diversification events reflected emergent ages (but not current sizes) of our pre-defined areas; areas with recent histories of sea-level inundation (e.g., Sumatra and Java) were characterized by higher levels of (recent) immigration than emigration; number of colonization events were greater between adjacent areas compared to more distant areas; and finally, an increase in diversification rates was evident during the Plio-Pleistocene, likely resulting from increased connection-disconnection events related to sea-level fluctuations.

While the importance of Indochina to the evolutionary history of SE Asian biodiversity has long been recognized (Myers et al., 2000; Lohman et al., 2011), Borneo, embedded in one of the world's most exceptionally diverse biodiversity hotspots (Sundaland; see Myers et al., 2000) is clearly also of extraordinary importance to the SE Asian biota, as shown here across several taxonomic groups through all time periods investigated. Much of Borneo's remarkable diversity, including remnant populations of conservation icons such as Sumatran rhinoceros (*Dicerorhinus sumatrensis*) and Bornean orangutan (*Pongo pygmaeus*), is under extreme and sustained pressure from anthropogenic habitat conversion (Miettinen et al., 2011) and unsustainable use (Koh & Sodhi, 2010; Wilcove et al., 2013).

These results conclusively drive home the importance of maintaining extensive high priority conservation areas in Borneo as envisaged under the current 'Heart of Borneo' agreement between Indonesia, Malaysia and Brunei. However, the agreement fails to adequately protect extensive areas of lowland rainforest, which harbor the highest levels of diversity. Losing further large areas of forested land to development in this region will result in the irreplaceable loss of the primary refuge area for the entire Sunda Shelf region.

Chapter 2

The biogeography of Sulawesi revisited: is there evidence for a vicariant origin of taxa on Wallace's "anomalous island"?

Introduction

> "We now come to the Island of Celebes, in many respects the most remarkable and interesting in the whole region, or perhaps on the globe, since no other island seems to present so many curious problems for solution."
>
> Wallace (1876)

The region of oceanic islands in the centre of the Indo-Australian Archipelago (IAA) comprising Sulawesi, the Moluccas and the Lesser Sunda Islands has been dubbed Wallacea by Dickerson (1928). The largely endemic fauna of this biodiversity hotspot (Myers et al., 2000) is sharply distinct from that of the continental islands on the adjacent shelves, the Asian (Sunda) Shelf with Borneo, Java and Sumatra in the West and the Australian (Sahul) Shelf with New Guinea in the East. The western and eastern limits of Wallacea coincide with Wallace's Line (Wallace, 1863) and Lydekker's Line (Lydekker, 1896), respectively, which were originally proposed as faunistic boundaries between the Asian and Australian fauna in recognition of the faunal breaks observed by the early naturalists (for a historical overview see Mayr, 1944; Simpson, 1977). In the North, Wallacea borders on the other large region of oceanic islands in the IAA, the Philippines, which was originally included in Wallacea by Dickerson (1928), but is now generally regarded as a distinct biogeographic entity (but see Cox, 2001; Michaux, 2010), as it has a largely unique fauna of its own (see e.g., Heaney et al., 2005) and is regarded as a biodiversity hotspot as well (Myers et al., 2000). The biogeography of the IAA in general and Wallacea in particular has been a major research issue ever since the seminal work by Wallace (Wallace, 1869). Recently, i.e., during the last 10-20 years, the widespread application of molecular phylogenetics to the region's taxa and an increasingly detailed understanding of its complex geology have provided a boost to respective studies (see review by Lohman et al., 2011). Today, most of the Philippine and Wallacean islands are generally regarded as true oceanic islands that have had no terrestrial connection to any surrounding area since their emergence (van Oosterzee, 1997).

The biogeography of Sulawesi

Sulawesi, the former Celebes, is the largest and possibly oldest island in Wallacea and as such harbours the most diverse fauna (see e.g., Whitten et al., 2002). Wallace's difficulties in the taxonomic assignment of some characteristic taxa of Sulawesi, such as the well-known babirusa (*Babyrousa babyrussa* Perry 1811), to living relatives within their family led him to regard it as comprising "remnants of an exceedingly ancient land" (Wallace, 1876) or simply an "anomalous island" (Wallace, 1880). Soon and perhaps not very surprisingly given geographic position, it had been realized that Sulawesi's fauna is predominantly of Asian origin (Sarasin & Sarasin, 1901; Wallace, 1910; de Beaufort, 1926; Stresemann, 1939; Whitmore, 1987; Holloway, 1990; Whitten et al., 2002). This notion is supported by most recent molecular

phylogenies involving taxa from across the Wallace line which have also identified (comparatively) recent Miocene to Pleistocene dispersal as the most likely mechanism for the origin of these taxa on Sulawesi (see e.g., Butlin et al., 1998; Evans et al., 1999, 2003a; Inger & Voris, 2001; Mercer & Roth, 2003; Alfaro et al., 2008; Merker et al., 2009). Nevertheless, the origin of some endemic taxa on Sulawesi is still controversial, particularly with regard to the role of vicariance and dispersal in bringing about current distribution patterns (Lohman et al. 2011). The complex geological history of Sulawesi offers the theoretical possibility for a vicariant origin of taxa both from across the Wallace Line, i.e., from Sundaland, and from New Guinea/Australia (see below for details), and several biogeographers or geologists have suggested respective scenarios (Burrett et al., 1991; Michaux, 1991, 1994, 1996; de Boer & Duffels, 1996; Moss & Wilson, 1998). Consequently, in several recent molecular phylogenetic studies estimating divergence times of Sulawesi taxa a vicariant origin has been suggested or at least considered (from Asia: Walton et al., 1997; Maekawa et al., 2001; Takehana et al., 2005; Clouse & Giribet, 2010; from Australia: Glaubrecht & von Rintelen, 2003; Sparks & Smith, 2004; Ruedas & Morales, 2005; von Rintelen & Glaubrecht, 2005). Recently, Michaux (2010) analysed the topologies from 33 molecular phylogenies comprising Sulawesi taxa, albeit without considering divergence time estimates, and suggested an vicariant origin of a considerably higher number of groups on Sulawesi, either across Wallace's Line (23 taxa) or from Australia (9 taxa).

The geology of Sulawesi

Geologically, the IAA is highly complex and the site of an on-going collision between two plates (Asia and Australia) which began in the Miocene involving numerous allochthonous continental fragments (terranes) (Hall, 2009a). Plate tectonic reconstructions of the geological evolution of the IAA through the Cenozoic have been constantly refined (Hall, 1996, 1997, 2001, 2002, 2011) but many uncertainties remain. The complex history is mirrored in the geology of Sulawesi, which is a composite island at the centre of this collision zone. For many years, Sulawesi has been divided into three major tectonic provinces, the (1) West Sulawesi Plutono-Volcanic Arc, the (2) East Sulawesi Ophiolite Belt, the (3) Central Sulawesi Metamorphic Belt, with two smaller continental fragments, Banggai-Sula in the north-east and the Tukang Besi Block in the southeast (e.g., Hamilton, 1979; Moss & Wilson, 1998; Fig. 2.1). This has led to the view that Sulawesi represents the collision between an active volcanic margin and multiple micro-continental fragments sliced from New Guinea which arrived at different times since the Early Miocene.

Recent work has considerably changed this simple picture (Hall, 2009a, 2011; Spakman & Hall, 2010). By the Late Cretaceous West Sulawesi was part of the Asian margin and became separated from it in the Eocene (c. 45 Mya) by rifting that led to formation of the

Makassar Straits (Fig. 2.1, red arrow #1). West Sulawesi was not the site of significant active magmatism but there was a volcanic arc in the North Arm of Sulawesi from the Eocene. In the Early Miocene (c. 23-20 Mya) a large promontory, the Sula Spur, which was the continuation of the Australian continental margin in New Guinea, collided with the North Sulawesi volcanic arc resulting in emplacement of ophiolites, and probably resulting in emergence of much of east and southeast Sulawesi. Subduction continued at the Java Trench and this subduction rolled back into the Banda oceanic embayment south of the Sula Spur from the Middle Miocene leading to extension of Sulawesi and the Sula Spur. It was this extension that formed the micro-continental fragments, such as Banggai-Sula and Tukang Besi, separated by young oceanic crust of the North and later South Banda Seas. This extensional fragmentation began about 15 Mya (Fig. 2.1, red arrow #2). The Banda volcanic arc formed as the subduction rolled back eastwards into the Banda embayment, and this collided with the southern continental margin of the embayment in Timor at about 4 Mya. The movie in Spakman & Hall (2010; http://searg.rhul.ac.uk/pubs/spakman_hall_2010_banda) shows this geological history which, although very different from the traditional view of slicing of fragments from New Guinea and their translation westwards as potential arks (cf. Michaux, 2010), did lead to fragmentation which could account for vicariant biogeographic hypotheses. It is still unclear at what time the collision of the Sula Spur with the North and West arms of Sulawesi led to the establishment of a (subaerial) connection between these and the fragments of the island derived from the Sula Spur (Fig. 2.1, red arrow #3), which is reflected in a wide time range for this event (10-20 Mya).

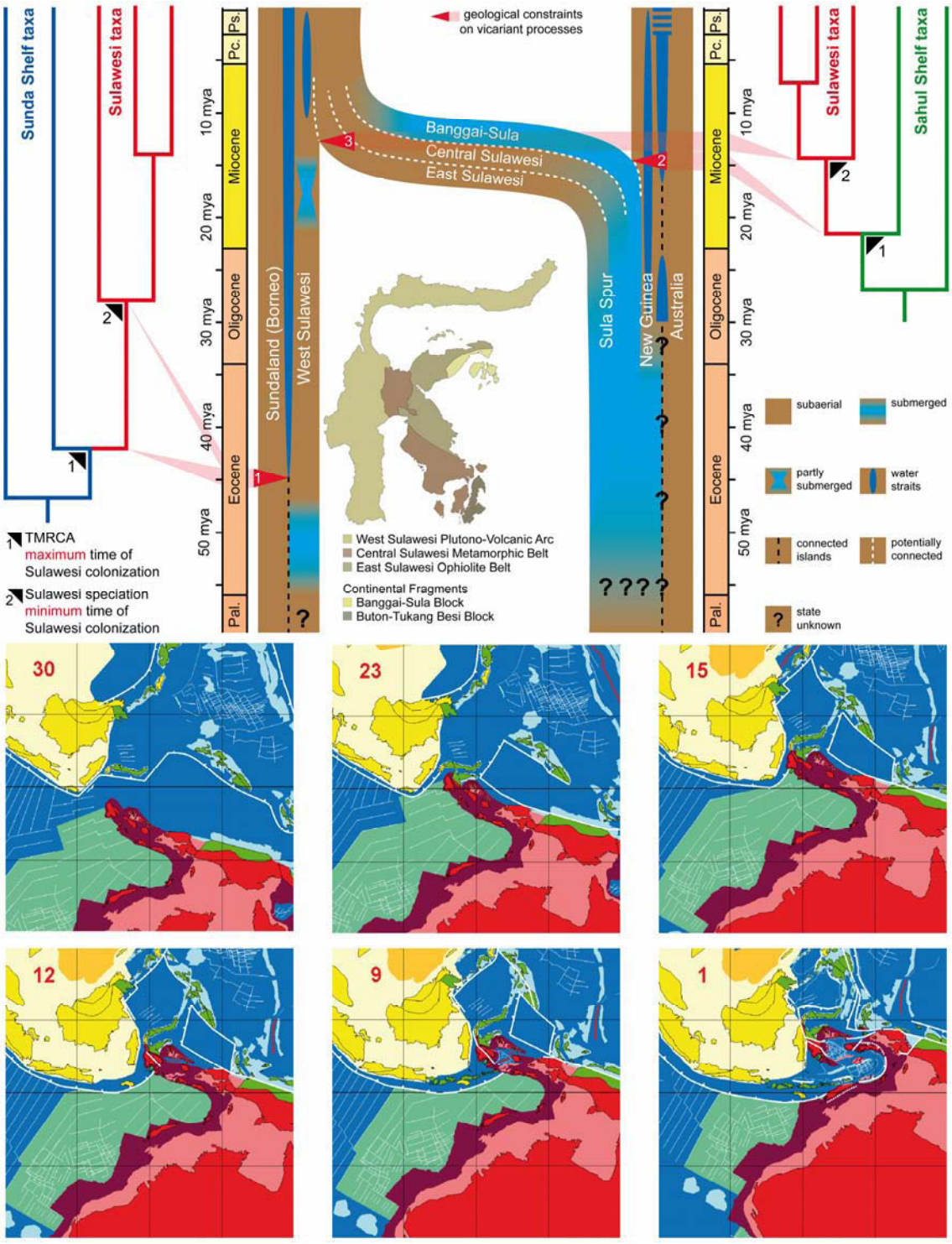


Figure 2.1. The tectonic history of Sulawesi and geological constraints on vicariant events. A. Schematic summary of the geographic connections and the timing of separation or collision of the different parts of Sulawesi shown in the sketch map of the island. Please compare text for details. Blue tinting indicates presumably submerged areas, see legend for details. Red arrows indicate area vicariance (1,2) or fusion (3). Based on Moss & Wilson (1998), Hall (2009b), and Spakman & Hall (2010). B, C. Hypothetical phylogenies of Sulawesi taxa B. derived from Asia and C. from Australia. The black triangles indicate (1) the split of the Sulawesi lineage from its sister taxon and (2) the first diversification event on Sulawesi. Under a vicariance scenario, the timing of these splits is constrained by the geological events indicated in panel A. D. Tectonic reconstructions of the Asia-Australia collision zone. Modified from Spakman & Hall (2010).

The importance of palaegeography for biogeographers

For biogeographers, the plate tectonic reconstructions of the IAA through the Cenozoic provide an indispensable framework for discussing the plausibility of dispersal and vicariance scenarios in the region. However, at least as important in this context is the past distribution of land and sea. If continental fragments are to act as 'rafts' for terrestrial or limnic fauna (dubbed 'tectonic dispersal' by Michaux, 2010), they must obviously have been subaerial throughout their existence as separate entities. While the formation of present-day Sulawesi is not due to the subsequent collision of micro-continental fragments with an Australian origin as previously believed (see above), the basic requirement of the continuous presence of land holds also true in the case of a single collision of the Sula Spur. For Sulawesi (and Wallacea in general), there is apparently little evidence for the relevant areas being continuously subaerial, with the best chance for the continuous existence of at least small amounts of land in Southwest Sulawesi (Fig. 2.1; see Hall, 2001, 2009a, but Moss & Wilson, 1998). If true, this would make any vicariance scenario at least by 'tectonic dispersal' to Sulawesi from the East a priori impossible. However, dynamic land connections, i.e., land is continuously present but fragmented and with shifting boundaries, which would require terrestrial or limnic organisms to migrate continuously in order to keep afloat, seem possible. The past emergence of smaller islands is notoriously difficult to infer, as characteristic signatures of larger land areas such as fluvial sediments will not be present, and the existence of land is largely inferred from negative evidence such as gaps in sediments of marine origin (Hall, 2009b). There remains a token of hope for biogeographers, thus, and it may be expected that future work is likely to lead to considerable revisions of paleogeographic maps of these critical regions (Hall, 2009b).

A comparative approach to Sulawesi's biogeography

Molecular phylogenetic data are now available for a wide range of Sulawesi taxa and most respective studies also discuss their biogeographic origin against a geological background. However, a comparative assessment of biogeographic patterns for Sulawesi taxa utilizing the temporal dimension, i.e., divergence time estimates, is still lacking. Here, we re-analyse published molecular data for (mostly non-flying) terrestrial or limnic taxa using standardized molecular clock analyses in order to test the compatibility of divergence time estimates with dispersal and vicariance hypotheses as proposed originally. We also explore the validity of our results with respect to the uncertainties inherent in molecular clock analyses and the vagaries of interpreting dated phylogenies in biogeographic terms.

Materials and Methods

Analysed studies

Altogether 27 individual datasets (alignments of individual genes) from 25 publications covering 20 different taxa were analysed (Table 2.1 and 2.2). In several of the original studies multiple markers were analysed, but not all single markers contained in these datasets could be included in the analyses as for some mtDNA and nuDNA datasets there were no external rates or calibration data available. Datasets containing multiple markers (see Tables 2.1 and 2.2) were analysed both jointly and separately, with the exception of the mite harvestmen dataset, which was only analysed jointly.

Table 2.1. Sulawesi taxa included in this study with a summary of divergence times and biogeographic hypotheses as originally published. For comparison, the hypotheses supported in this study and the maximum mean age of the time to the most recent common ancestor of a Sulawesi taxon (TMRCA) are also provided, see Tables 2.2 and S1 for details). If Sulawesi was colonized multiple times by a taxon (see respective column) only the age of the oldest lineage on the island is given here, see Fig. 2.2-4 and Table S1 for all divergence time estimates. The individual genes contained in multi-gene datasets were analyzed separately here (see Materials and Methods), the numbers of the respective dataset correspond to those used in Fig. 2.2-5 and Tables 2.2 and 2.3. D – dispersal; DV – mixed dispersal/vicariance model; V – in column *source*: vicariance hypothesis proposed or considered; in column *this study*: vicariance hypothesis compatible with geological constraints as shown in Fig. 2.1; V? – vicariance hypothesis not supported if TMRCA estimates based on 2nd substitution rate are used. Grey shading indicates major conflict between data or hypotheses from the source publication with the results from this study.

Taxon[1]	**Biogeographic hypothesis**		**Age of Sulawesi taxon (TMRCA, My)**[2]		**Colonizations of Sulawesi**	**Genes analyzed**	**Data-sets**
	source	**this study**	**Source**	**this study**			
Mite harvestmen[a]	V (Asia)	V (Asia)	c. 50	60.8 (43.9, 77.4)	2	concatenated dataset (16S, 18S, COI, H3, H4, 28S)	1
Water beetles[b]	D (Asia)	D (Asia)	c. 43	37.5 (-) | 24.9 (-)	2	Cyt b	2
				15.7 (-) | 2.1 (-)		COI	9
				3.8 (-)		16S	22
				10.2 (-)		Cyt b + COI + 16S	2+9+22
Freshwater snails[c]	V (Australia)[†]	V (Australia)	c. 16	30.5 (20.0, 42.0)	1	16S	3
Freshwater crabs[d]	-	D (Asia)	-	26.0 (12.1, 42.4)	1	16S	4
Shrews[e,f]	D (Asia)	D (Asia)	Pliocene	25.9 (19.3, 32.7) | 10.5 (7.9, 13.2)	2	Cyt b	5
			Pleistocene	21.4 (16.9, 26.1) | 8.7 (7.0, 10.5)			6
Macaques[g,h]	D (Asia)	D (Asia)	3-4.6	19.5 (12.6, 27.2) | 8.1 (5.4, 11.3)	2-3(**2**)*[§]	NADH	7
			Pleistocene	3.7 (2.5, 5.0) | 1.5 (1.0, 2.0)		12S	24
				18.6 (10.8, 29.0) | 7.6 (4.4, 11.8)		NADH + 12S	7+24
Phalangerids[i]	DV (Australia)	V (Australia)?	21.1-23.3	18.8 (14.1, 24.5) | 7.6 (5.6, 10.0)	1	16S	8
Sailfin silversides[k]	V (Australia)	V (Australia)?	25-30	15.2 (8.1, 23.4) | 12.5 (6.4, 19.3)	1	ND5	10
Cockroaches[l]	V (Asia)	D (Asia)	24-56	14.0 (10.9, 17.3)	2	COII	11
Squirrels[m]	D (Asia)	D (Asia)	11	12.6 (-)	1	16S	12
				11.8 (-)		12S	13
				11.1 (-)		16S + 12S	12+13
Megapodes[n]	-	D (Australia)	-	9.9 (6.6, 13.8)	1	ND2	14
Fanged frogs[o]	D (Asia)	D (Asia | Philippines)	Pliocene	9.6 (6.9, 12.5) | 5.7 (3.8, 7.3)	2-3(**2**)**	16S	15
				9.2 (-) | 5.6 (-)		12S	16
				9.4 (6.91, 11.7) | 5.7 (4.2, 7.2)		16S + 12S	15+16
Tarsiers[p]	D (Asia)[$]	D (Asia)	11[‡]	8.3 (3.9, 16.7) | 3.9 (1.6, 9.4)	1	Cyt b	17
Grasshoppers[q]	V (Asia)	D (Asia)	7-14	7.9 (2.1, 8.0)	1	COI	18
Ricefishes[r]	V (Asia)	D (Asia)	29-32	7.1 (4.4, 10.2) | 5.7 (3.6, 8.1)	1	16S	19
				3.8 (2.6, 5.1) | 3.0 (2.0, 4.0)		12S	23
				4.9 (3.6, 6.5) | 3.9 (2.8, 5.1)		16S + 12S	19+23
Bovids[s]	-	D (Asia)	2.23	4.3 (3.2, 5.5) | 1.8 (1.3, 2.3)	1***	Cyt b	20

Table continued

Taxon[1]	Biogeographic hypothesis		Age of Sulawesi taxon (TMRCA, My)[2]		Colonizations of Sulawesi	Genes analyzed	Data-sets
	source	this study	Source	this study			
Toads[t]	D (Asia)	D (Asia)	Pleistocene	4.2 (2.6, 5.9) | 2.6 (1.5, 3.6)	1	12S	21
Water snakes[u]	D (Asia)	D (Asia)	0.5-4.5	2.1 (1.1, 3.4)	1	Cyt b	25
Freshwater bivalves[v]	D (Asia)	D (Asia)	-	1.2 (-) | 1.4 (0.9, 1.9)	3	COI	26
Pigs[w]	D (Asia)	D (Asia)	Pleistocene[z]	0.9 (-)	3***	CR	27

[1] Dataset: a) Clouse & Giribet (2010); b) Balke et al. (2004); c) Köhler & Glaubrecht (2010); d) Schubart & Ng (2008); e) Esselstyn et al. (2009); f) Ruedi et al. (1998); g) Evans et al. (1999); h) Evans et al. (2003b); i) Ruedas & Morales (2005); k) Sparks & Smith (2004); l) Maekawa et al. (2001); m) Mercer & Roth (2003); n) Birks & Edwards (2002); o) Evans et al. (2003a); p) Merker et al. (2009); q) Walton et al. (1997); r) Takehana et al. (2005); s) combination of Kikkawa et al. (1997), Hassanin & Douzery (1999), Schreiber et al. (1999) and Hassanin & Ropiquet (2004); t) Evans et al. (2003c); u) Alfaro et al. (2008); v) von Rintelen & Glaubrecht (2006); w) Larson et al. (2007). [2] Age of Sulawesi taxon: TMRCA of oldest lineage; 95% confidence intervals are given in brackets. Inferred divergence times separated by vertical lines if different rates were used. [z] No divergence time estimate attempted or discussed in original study, minimum age based on fossil record of close relative in Europe. * Topology conflict in the reduced concatenated dataset, but no support in basal splits. The number of colonizations inferred in the source publication is highlighted in bold type. [§] Based on evidence from several nuclear markers, Sulawesi macaques may form a monophyletic group (Tosi et al., 2003; Evans et al., 2010). ** Differs between datasets in this study (topology conflict among both rates used in 12S, but no support for three lineages; 2 colonizations in 16S and in the concatenated dataset). The number of colonizations inferred in the source publication is highlighted in bold type. *** Topology of tree in our study would suggest an origin of the Asian clade on Sulawesi, but has been ignored here because of the complete lack of nodal support (see Fig. S22, S27). [$] Transitory land bridge also considered in original study. [‡] A priori assumption in original study, not based on divergence time estimate. [†] = Köhler & Glaubrecht (2010) do not discuss the biogeography of the Sulawesi lineage included in their study, but vicariance has been explicitly assumed by Glaubrecht & von Rintelen (2003) and von Rintelen & Glaubrecht (2005) based on molecular phylogenetic data without divergence time estimates.

Molecular clock analyses

Sequences were obtained from GenBank or were copied from the published alignment; base ambiguity codes (R, Y, M, K, W, S) were replaced by N in order to avoid an overestimation of the number of haplotypes due to potential sequencing errors. Noncoding sequences were aligned using MAFFT (Katoh & Toh, 2008); default settings: gap open = 1.53, gap extension = 0.123, perform FFTS = localpair) and corrected by eye. All datasets were reduced to unique haplotypes (sequences containing more than 30% of N were deleted *a priori*) using DAMBE v. 5.1.1 (Xia & Xie, 2001) and were subsequently tested for nucleotide substitution saturation using the test by Xia & Xie (2001) as implemented in DAMBE including the estimation of the proportion of invariant sites (Pinv; Pinv tree-building algorithms: FastME, default settings). For the included datasets, tests revealed no significant saturation ($p > 0.05$), except for the assumption of an (unlikely) completely asymmetrical tree in some datasets (2, 5, 9, 10, 14, 22-24). Substitution models for Bayesian inference (BI) analyses and molecular clock analyses were estimated using the AIC in jModelTest (Posada, 2008); 3 substitution schemes, i.e., 24 models) (Table 2.2).

Molecular clock analyses can be performed using either strict or relaxed clock conditions, i.e., under the assumption of homogeneous and heterogeneous substitution rates among branches, respectively. As a comparative approach, two different analysis methods have been conducted to test for the applicability of a strict clock. First, BI analyses were performed using MrBayes v. 3.1.2 (Ronquist & Huelsenbeck, 2003; parameters: ngen = 500,000-5,000,000, samplefreq = 10-100 (depending on ngen size), nchains = 4, burnin value = ngen-x = 15,000 (samples used for analyses)); a second BI run was performed ($ngen_{strict} = ngen_{default}$) for all datasets including a strict clock model (both sump burnin outputs were saved). Secondly, strict and relaxed lognormal molecular clock analyses were performed using BEAST v. 1.5.3 (Drummond & Rambaut, 2007; parameters used: Speciation: Yule process, ngen = 20,000,000, log = 400, burnin value = 35,001, calibration point setting for datasets 1, 10, 14, 25: normal distribution, i.e., this prior – in comparison to a lognormal distribution – allows a younger age of the respective fossil used for the calibration). For two datasets (mite harvestmen and freshwater snails), strict and relaxed clock analyses were additionally performed under a birth-death process to assess the potential influence of extinction. Effective sampling size (ESS) values never dropped below 200 (Table 2.2). Specific clock rates and calibration points, respectively, are given in Table 2.2. Whenever possible, a taxon and gene specific external rate or calibration point was used, but for some data sets this was not possible and 'general' substitution rates were used (see Table 2.2). Unspecific rates were taken from the next relative of the target taxon for which a rate was available, at best from the same family, but more frequently from the same major clade, e.g., amphibians or insects (see Table 2.2 for details).

Bayes factor analyses for all datasets were conducted in order to test whether a strict clock could be accepted *a priori* by using harmonic mean values (the total value of run1 and

run2 from both default and strict-enforced clock runs obtained from the sump output in MrBayes) and calculating twice the difference (DHM; delta harmonic mean). For the BEAST runs, a Bayes factor analysis ($\log_{10}$ Bayes factors) was conducted as a *post-hoc* test using tree likelihoods of both strict and relaxed lognormal clock analyses in Tracer v. 1.5 (Rambaut & Drummond, 2007); 1000 bootstrap replicates) to assess support for a strict *vs.* a relaxed clock. Negative Bayes factor values are linked to a higher probability of the null hypothesis (i.e., strict-clock enforced runs in MrBayes and strict clock models in BEAST) and vice versa (Kass & Raftery, 1995; see also Suchard et al., 2001). Kass & Raftery (1995) also suggest thresholds for deciding in favour of or against the null hypothesis: 0-3 (positive support), 3-6 (strong support), and > 6 (decisive support; see also Table 2.2).

Coefficients of variation (COV; obtained from the log files imported in Tracer) of the relaxed clock BEAST analyses were also considered as this parameter represents an indication for rate heterogeneity across the tree. Small COV values close to zero denote a clock-like evolution among lineages, whereas larger values indicate a higher degree of rate heterogeneity and may give a higher plausibility to perform relaxed clock analyses (e.g., Drummond et al., 2006).

Table 2.2. Datasets studied including test results and parameters for the molecular clock analysis. Estimated clock (clock rate and ucld mean) rates are per My. Abbreviations used: Pinv – proportion of invariant sites; na – not applicable; BF(DHM, MrBayes) – Bayes factors of the MrBayes analyses using the difference of the harmonic means (see also text); log_{10}BF – log_{10}Bayes factors comparing relaxed and strict treeLikelihoods of BEAST analyses in Tracer. According to Kass & Raftery (1995): *** = decisive support for null hypothesis (relaxed clock), BF values > 6; ** = strong support for null hypothesis, BF values 3-6; * = positive support for null hypothesis, BF values 0-3.

Taxon	Dataset(s)	Gene	Pinv	BF MrBayes (DHM)	Strict clock rejected? DHM	log_{10}BF BEAST | COV (coefficient of variation)	Substitution model	Clock rate (rate 1 | rate 2) or calibration (My +/- SD) used → source[1]	Estimated clock rate | ESS
Mite harvestmen	1	multiple*	na	na	na	158.680*** | 0.581	GTR+I+G	425 +/- 1 → 11	0.001 | 21
Water beetles[#]	2	Cyt b	0.088	-6.28	no	6.879*** | 0.215	HKY+I+G	0.023 | 0.0343 → 1 / 12	-
	9	COI	0.263	30.28***	yes	8.433*** | 0.313	GTR+I+G	0.023 | 0.17212 → 1 / 12	-
	22	16S	0.365	45.26***	yes	7.020*** | 0.391	GTR+I+G	0.023 → 1	-
Freshwater snails	3	16S	0.428	21.94***	yes	7.155*** | 0.385	GTR+I+G	0.01 → 8	-
Freshwater crabs	4	16S	0.598	2.32*	yes	1.988* | 0.729	GTR+I+G	0.0088 → 2	-
Shrews	5	Cyt b	0.561	-7.56	no	0.219* | 0.113	GTR+I+G	0.00562 | 0.01385 → 5	-
	6	Cyt b	0.554	-30.28	no	0.518* | 0.008	GTR+I+G	0.00562 | 0.01385 → 5	-
Macaques[#]	7	NADH	0.601	179.74***	yes	41.727*** | 0.840	GTR+I+G	0.00562 | 0.01385 → 5	-
	24	12S	0.683	-94.10	no	1.188* | 0.366	HKY+I+G	0.00562 | 0.01385 → 5	-
Phalangerids	8	16S	0.686	168.40***	yes	0.438* | 0.204	GTR+G	0.00562 | 0.01385 → 5	-
Sailfin silversides	10	ND5	0.464	5.60**	yes	1.507* | 0.274	GTR+I+G	0.025 | 0.031 → 4	-
Cockroaches	11	COII	0.068	1.00*	yes	3.726** | 0.223	GTR+I+G	0.023 → 1	-
Squirrels[#]	12	16S	0.012	59.08***	yes	14.613*** | 0.369	GTR+I+G	36 +/- 1 → 3	0.0110 | 1251
	13	12S	0.019	114.92***	yes	22.011*** | 0.390	GTR+I+G	36 +/- 1 → 3	0.0132 | 541
Megapodes	14	ND2	0.465	47.52***	yes	11.466*** | 0.615	GTR+I+G	0.0179 → 6	-
Fanged frogs[#]	15	16S	0.068	168.40***	yes	36.363*** | 0.447	GTR+I+G	0.0148 | 0.0245 → 7	-
	16	12S	0.004	74.84***	yes	19.187*** | 0.447	HKY+I+G	0.0148 | 0.0245 → 7	-
Tarsiers	17	Cyt b	0.103	-15.12	no	1.879* | 2.621	GTR+G	0.00562 | 0.01385 → 5	-
Grasshoppers	18	COI	0.738	-9.92	no	8.401*** | 1.739	GTR+I+G	0.023 → 1	-
Ricefishes[#]	19	16S	0.004	18.82***	yes	5.364** | 0.600	GTR+I+G	0.025 | 0.031 → 4	-
	23	12S	0.492	-28.38	no	0.977* | 0.280	GTR+I+G	0.025 | 0.031 → 4	-
Bovids	20	Cyt b	0.064	-14.36	no	17.930*** | 0.266	GTR+I+G	0.00562 | 0.01385 → 5	-
Toads	21	12S	0.711	-44.14	no	0.080* | 0.341	GTR+I+G	0.0148 | 0.0245 → 7	-
Water snakes	25	Cyt b	0.446	24.70***	yes	4.247** | 0.272	GTR+I+G	36 +/- 19 → 9	0.0239 | 1096
Freshwater bivalves	26	COI	0.042	-135.98	no	5.595** | 0.740	HKY+I+G	0.0148 | 0.016 → 10	-
Pigs	27	CR	0.216	-35.18	no	19.803*** | 1.183	GTR+I+G	0.05 → 13	-

* Concatenated multiple gene dataset, see Table 2.1. [#] Also analyzed as concatenated dataset with same priors (clock model, substitution model, clock rate). In water beetles (datasets 2, 9 and 22) the concatenated dataset was analyzed using the priors: GTR+I+G, relaxed clock, clock rate: 0.023. [1]Rates (clock.rate and ucld.mean set in BEAUTi) and calibration points were derived from these sources: 1) Brower (1994), genetic distance in butterflies; 2) Schubart et al. (1998), 16S, freshwater crabs; 3) Mercer & Roth (2003), used with SD +/- 1; 4) Echelle et al. (2005), ND2, cyprinodont fishes; 5) Pesole et al. (1999), see also Esselstyn et al. (2009), mammalian mtDNA; 6) Pereira & Baker (2006), ND2, galliform birds; 7) Crawford (2003), ND2, leptodactylid frogs; 8) Köhler & Glaubrecht (2010), 16S, freshwater snails; 9) Alfaro et al. (2008), calibration point from fossil data; 10) Wilke et al. (2009), COI, mean +/- SD from 2 different gastropod datasets; 11) Clouse & Giribet (2010), calibration point from fossil data corresponds to root height; 12) Pons et al. (2010), cyt b and COI, beetles; 13) Tamura & Nei (1993), control region, primates.

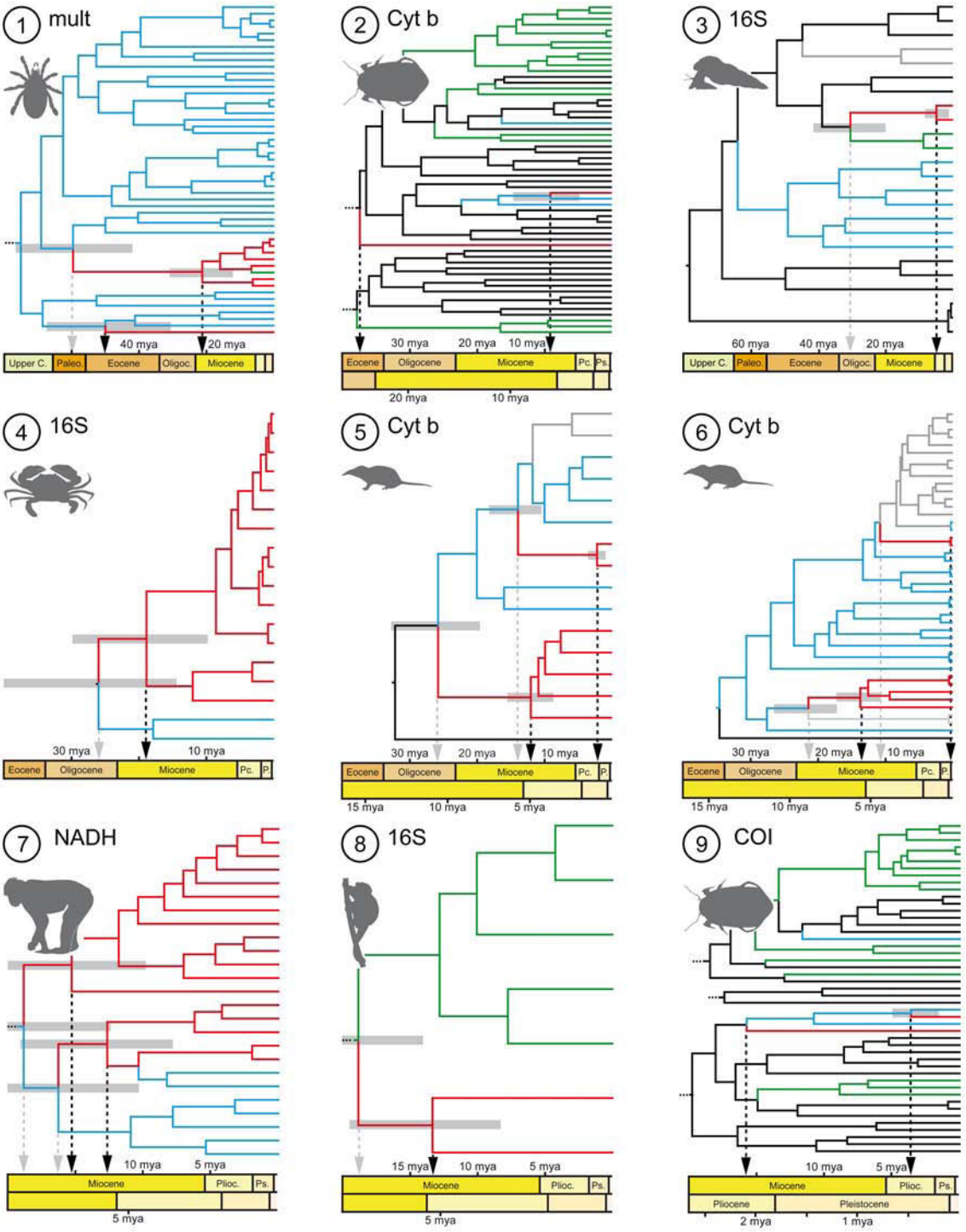


Figure 2.2. Calibrated phylogenies of Sulawesi taxa (1) arranged in approximately chronological order with decreasing age. Numbers refer to the respective dataset (Table 2.1). Diversification time is indicated by scale(s) beneath trees for each rate/calibration. Red – Sulawesi taxa; blue – Asian taxa; green – Australian (including New Guinea) taxa; grey = Philippines, Moluccas, Lesser Sunda Islands East of Bali. Arrows indicate the timing of Sulawesi diversification (black) and the split from the MRCA (grey). Horizontal bars at nodes show the 95% confidence interval of diversification time.

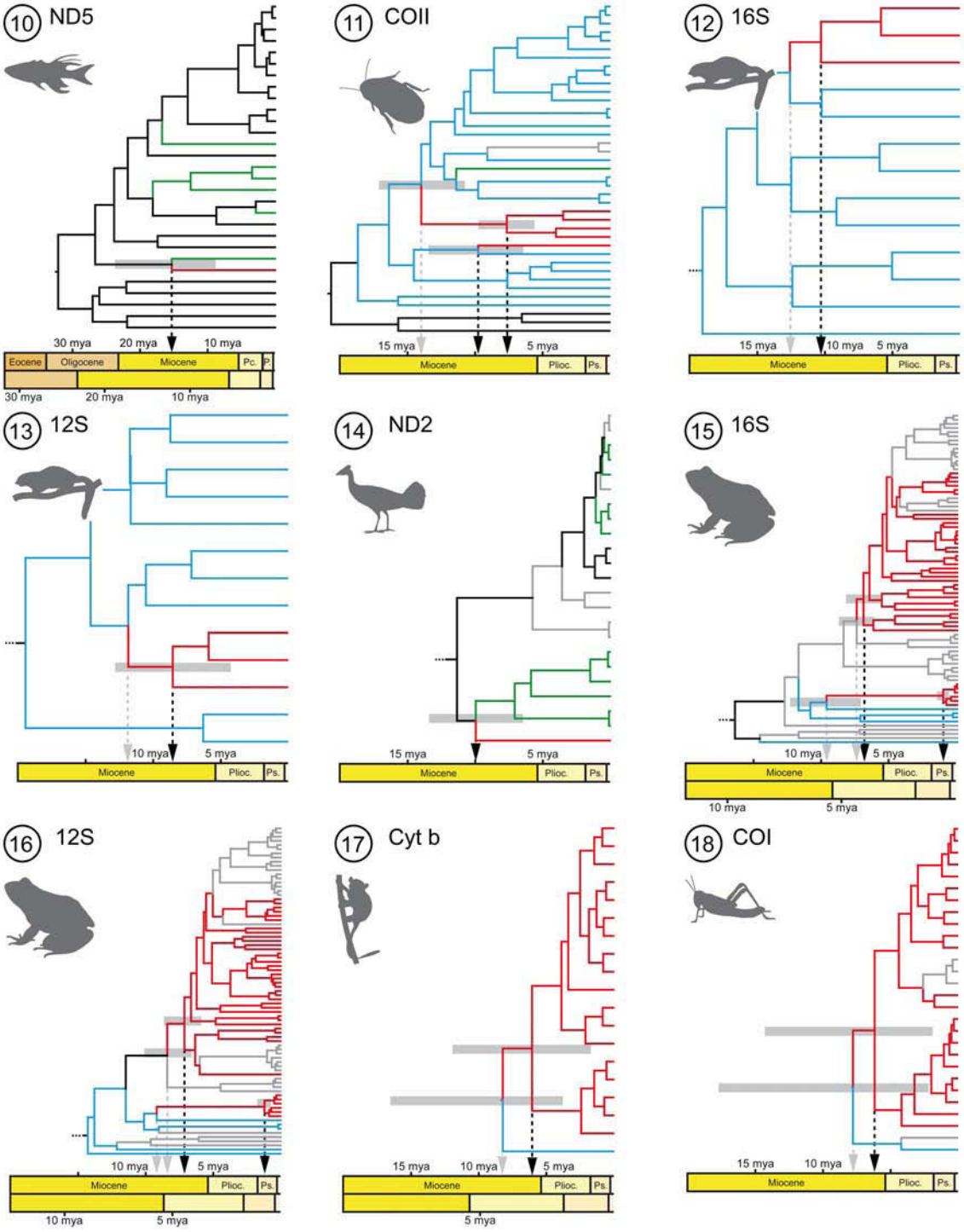


Figure 2.3. Calibrated phylogenies of Sulawesi taxa (2) arranged in approximately chronological order with decreasing age. Numbers refer to the respective dataset (Table 2.1). Diversification time is indicated by scale(s) beneath trees for each rate/calibration. Red – Sulawesi taxa; blue – Asian taxa; green – Australian (including New Guinea) taxa.

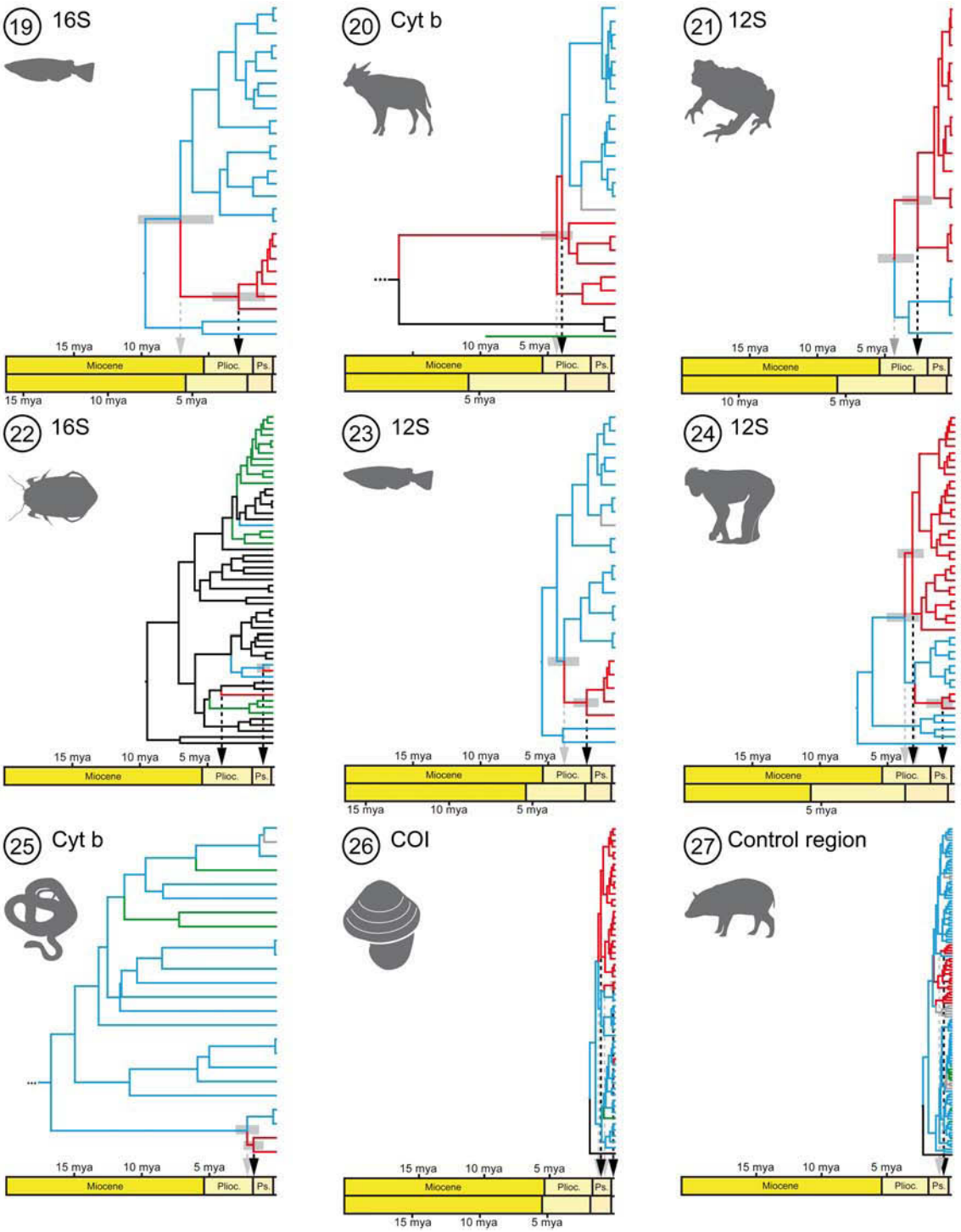


Figure 2.4. Calibrated phylogenies of Sulawesi taxa (3) arranged in approximately chronological order with decreasing age. Numbers refer to the respective dataset (Table 2.1). Diversification time is indicated by scale(s) beneath trees for each rate/calibration. Red – Sulawesi taxa; blue – Asian taxa; green – Australian (including New Guinea) taxa.

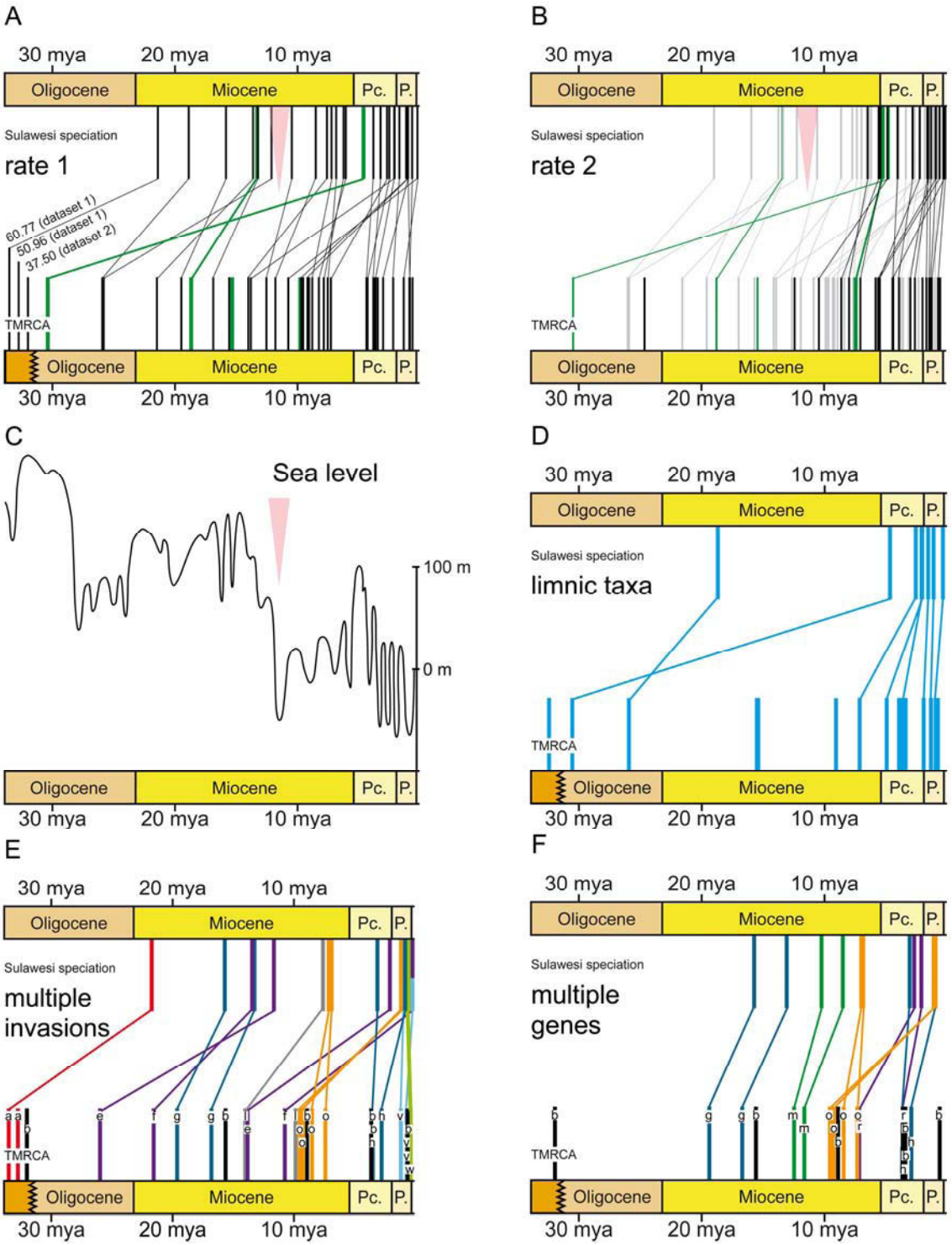


Figure 2.5. Summary of the geographic and temporal origin of Sulawesi taxa. The bars on the geological scales in each panel indicate the timing of (bottom) the split of the Sulawesi lineage from its sister taxon and of (top) the first diversification event on Sulawesi. See Table 2.1 for reference. A, B. Distribution of mean age, general patterns. Green bars indicate Sahul Shelf relationships and black bars Sunda Shelf relationships of Sulawesi taxa (sister group); pink arrows show sea-level lowstands (see panel C for details). A. Based on substitution rate 1 only; B. Based on substitution rate 2 only (if applicable; grey bars show rate 1 pattern, see panel A). C. Sea-level fluctuations in Southeast Asia, redrawn from Mercer & Roth (2003) and Haq et al. (1987); the red arrows denotes the onset of cyclic sea-level changes enabling Sulawesi colonization. D. Distribution of mean age of limnic taxa. In panels E and F, taxa are colour-coded and numbers on bars refer to the respective taxon (Table 2.1). E. Distribution of mean age of datasets revealing multiple colonizations of Sulawesi. F. Distribution of mean age patterns when the same substitution rate is applied to different mitochondrial markers.

Results

In the Bayes factor analyses using MrBayes, a strict clock was rejected for 15 out of 27 datasets (Table 2.2). The Beast analyses comparing a strict to a relaxed clock yielded always positive values, i.e., a strict clock was rejected for all datasets (Table 2.2). No clear link between Bayes factor support thresholds and COV values was found, though.

Node ages inferred by a relaxed clock differed by less than 10% from those estimated in the strict clock analyses in 56% (n = 15) of the datasets, with 22% (n = 6) differing by more than 20%. The tarsiers (dataset 17) are the only group showing a striking difference (> 50%) between the TMRCAs obtained from the two clock analyses, with a strong decrease of the TMRCA from 26.7 (rate 1) and 11.2 (rate 2) Mya (strict) to 8.3 and 3.9 Mya, respectively (relaxed). The following results are exclusively based on the relaxed clock analyses.

The topology of the individual calibrated phylogenies for the 20 taxa (27 datasets) analysed reveals an Asian (mainland and/or Larger Sunda Isles) origin for 15 taxa, a Philippine origin for one taxon, and an Australian (or New Guinea) origin for four taxa (Fig. 2.2-4, Table 2.1). In this respect, different datasets (= different genes) for the same taxon (e.g., for water beetles, datasets 2, 9, 22 or squirrels, datasets 12, 13) did not reveal any differences if node support is taken into consideration (Fig. S1-S27, Supplementary Information). The term origin is used in a loose sense here also for cases where Sulawesi taxa are merely sister group to an Asian or Australian lineage as, e.g., the pachychilid snails (Fig. 2.2, dataset 3), which is no evidence for an origin of the Sulawesi taxon from the respective area. Given the overall topologies of the trees (see also Supplementary Information), the origin ascribed here to a Sulawesi taxon seems nevertheless valid, if read as a metaphor for closest relatives being in Asia or Australia.

The analysis of concatenated multigene datasets for the respective taxa revealed no topological conflicts to the single gene analyses if node support is considered (Fig. S1-S27, Supplementary Information). For a few taxa with an origin in Asia, multiple independent colonizations of Sulawesi have been shown, e.g., for shrews or fanged frogs (Table 2.1, Fig. 2.2-5). The topologies reconstructed in our study generally match the phylogenetic patterns described in the source publications (see Table 2.1), at least with respect to the position of the Sulawesi taxa. Topological conflicts between our results and the original publications or between different datasets here had either no effect on the conclusions (see below, discussion) or were not supported at all, as in two of the three single gene water beetle datasets (Bayesian posterior probabilities below 0.5).

The estimation of divergence times, which are based on the results from the relaxed clock analyses, reveals a wide range from the Paleocene to the Pleistocene for the time to the most recent common ancestor (TMRCA; stem age, Crisp et al., 2011) of Sulawesi taxa, i.e., the

split from their sister group outside of Sulawesi (Fig. 2.5A, Table 2.1, Table S1). Except for two taxa, the mite harvestmen and water beetles (Fig. 2.2, datasets 1 and 2), the oldest TMRCA of any Sulawesi taxon is the early Oligocene, though, which narrows the range by c. 25 My. The time range estimated for the earliest speciation event on Sulawesi (crown age of Sulawesi lineages; Crisp et al., 2011) is less wide, the oldest split being in the early Miocene (c. 20 Mya). For several datasets the TMRCA was also estimated using an alternative substitution rate. In most cases the results of both analyses differed considerably (Table 2.1, Fig. 2.2-4, 5B).

The additional application of a birth-death (BD) process model to the mite harvestmen and pachychilid snail datasets revealed a decrease in node age for the older lineage in mite harvestmen by nine My to 51.8 Mya in comparison to the TMRCA estimate of 60.8 Mya under a Yule process, but for the younger lineage the TRMCA estimate increased to 54.0 Mya (BD) from 50.96 (Yule). For the pachychilid snails, the difference between TMRCA estimates under the two models was negligible (Yule: 30.5 *vs.* BD: 29.9 Mya).

Discussion

The origin of Sulawesi taxa

The divergence time estimates for Sulawesi taxa can be used to test vicariance hypotheses by contrasting the TMRCA with the temporal constraints on vicariance events derived from geology (Fig. 2.1). The TMRCA for a Sulawesi taxon must predate (but compare discussion below) or coincide with a vicariant event such as the opening of the Makassar Strait or the separation of the relevant areas of the Sula Spur from New Guinea (compare Fig. 2.1, red arrows labelled 1 and 2), otherwise a vicariance hypothesis is falsified (but see below for caveats). For the purpose of this discussion, the oldest time estimate has always been used, even though there is some evidence that the alternative rate may be the better choice, see discussion on the reliability on the molecular clock analyses below. Similarly, in case of multiple colonizations the older TMRCA was used (see Fig 5 D). Confidence intervals of divergence time estimates have been considered in all cases here, i.e., the interpretation is not only based on the mean TMRCA (compare Table 2.1 and Fig. 2.2-4). By applying these constraints, a different pattern emerges for taxa of Asian and Australian origin, respectively.

All post-Eocene TMRCA estimates of Sulawesi taxa postdate the opening of the Makassar Strait c. 45 Mya (see Introduction) and are thus not compatible with the assumption of a vicariance hypothesis for an origin of Sulawesi taxa from the West (Fig. 2.5A, Table 2.1). Three of the four studies on Sulawesi taxa of Asian origin that have considered a vicariance hypothesis involving the opening of the Makassar Strait have also found a post-Eocene TMRCA for the respective taxon (grasshoppers, Walton et al., 1997; ricefishes, Takehana et al., 2005; cockroaches, Maekawa et al., 2001), and support for a vicariance scenario was solely based on a

somewhat relaxed interpretation of the timeframe for the separation of Borneo and Sulawesi. Trans-Makassar Strait vicariance is only supported for one of the four taxa for whom a respective hypothesis has been suggested, viz. mite harvestmen (Clouse & Giribet, 2010), as the TMRCA estimates of c. 51 and 61 Mya found in our analyses for the two lineages on Sulawesi fit the geological timeframe (Fig. 2.2, Table 2.1). However, poor node support and low ESS values were found in both the strict and relaxed clock analyses, which might indicate general issues with the dataset as analysed.

Vicariant 'tectonic dispersal' hypotheses to account for the origin of taxa from the East on the Sula Spur have been proposed or at least discussed for three of the four Sulawesi taxa with Australian affinities considered here (pachychilid snails, Glaubrecht & von Rintelen, 2003; von Rintelen & Glaubrecht, 2005; Köhler & Glaubrecht, 2010; sailfin silversides, Sparks & Smith, 2004; phalangerids, Ruedas & Morales, 2005). The TMRCA for all three taxa either clearly pre-dates (pachychilid snails) or rather closely matches (sailfin silversides, phalangerids) the beginning separation of the Sula Spur parts that collided and remained with Sulawesi from New Guinea-Australia c. 15 Mya (compare Fig. 2.1, red arrow #2, Fig. 2.2-4; Table 2.1). For the latter two taxa this only holds true for one of the two substitution rates used for estimating divergence times, though, as the second rate yields a younger TMRCA in both cases (if confidence intervals are considered, the divergence time of sailfin silversides matches the timing of Sula-Australia split also assuming the faster rate). The TMRCA of the fourth Sulawesi taxon with closest relatives in Australia, the megapodes, clearly postdates the relevant vicariant events. However, the Sulawesi megapodes are probably the best dispersers among these taxa, since they are capable of flying at least short distances, depending on the taxon (Dekker, 2007).

In summary, the divergence time estimates derived by application of a standardized molecular clock approach are not compatible with a vicariance hypothesis in 80% of the taxa analysed and suggest dispersal as the predominant process in the origin of Sulawesi taxa. This pattern becomes even more pronounced if the faster of two external rates used for some datasets should turn out to be more appropriate for the respective taxon (Fig. 2.5B, here called rate 2; compare also discussion below). However, our data are *prima facie* also compatible with vicariance hypotheses in 20% of the analysed taxa. This proportion is 75% for taxa of Australian origin/affinities, but this can hardly be regarded as representative given the low number of respective datasets included.

Vicariance hypotheses involving Sulawesi taxa are at odds with the current paleogeographic reconstructions of Wallacea, i.e., the past distribution of land and sea during the Cenozoic. Recent updates of the respective models suggest that both in West Sulawesi and on the Sula Spur larger areas may have been subaerial since the Eocene than previously assumed (compare Hall 2009b and Lohman et al., 2011). There is as yet no evidence for the existence of a subaerial link between the Sula Spur and Australia-New Guinea during the

critical time period from the Oligocene to the early Miocene, though. Such a link is an essential assumption for matching a vicariance hypothesis with the respective dated phylogenies e.g., for the pachychilid snails. However, as emphasized above, reconstructions of the past distribution of land based on geological data tend to be biased towards detecting larger areas of land (Hall 2009b), and it should also be noted that the intervals of 5 or 10 My for which reconstructions are available for the last 40 My are rather coarse. We thus suggest not to reject a vicariance hypothesis for explaining the origin of Sulawesi taxa merely because it is not consistent with paleogeographic reconstructions, which are subject to at least as much uncertainty as the biological data. It seems fair to assume 'bottlenecks' in the area of available land both in Southwest Sulawesi and on the Sula Spur, though, under a vicariance scenario. Interestingly, the two taxa where there seems to be least temporal uncertainty about the compatibility with a vicariance scenario, the mite harvestmen and pachychilid snails, show an extreme time lag between their respective TMRCA and the first speciation event on Sulawesi (Fig. 2.5A). This pattern matches the assumption of the respective lineage persisting on very little land during the Miocene both in Southwest Sulawesi at around 20 Mya and on the Sula Spur until about 20 Mya, which would presumably constrain the potential for speciation in both taxa by limiting population size, habitat diversity, and the possibility for allopatry. However, a long time lag between the TMRCA and speciation on Sulawesi might also be, e.g., caused by insufficient taxonomic coverage or extinction. A role of the latter is possibly supported by the lower TMRCA found for the older Sulawesi lineage of mite harvestmen assuming a birth-death process, which also considers extinction in node age estimation. Nevertheless, a sampling bias in terms of taxa and area coverage is quite likely as well for a number of the invertebrates considered here, given the low number of sample sites on Sulawesi, but also other areas in some cases, e.g., the mite harvestmen (compare Clouse & Giribet, 2010: Fig. 2.1).

The earliest speciation event on Sulawesi in any taxon included in our study dates from the early Miocene c. 20 Mya (Fig. 2.5A). This might be linked to two geological/ climatological events: (1) the availability of more land on Sulawesi from about that time as a consequence of the collision of the Sula Spur with the North and West arms of the island (Hall 2009b; Spakman & Hall, 2010; compare Fig. 2.1) and (2) the climate-induced start of periodic sea level lowstands in the Mid-Miocene with inverse peaks in the late Miocene and Pliocene (see Haq et al., 1987; de Graciansky et al., 1998; Fig. 2.5C: pink arrow). Both factors would not only increase the potential of speciation on Sulawesi by providing e.g., opportunities for creating new ecological niches or allopatric differentiation, they would also increase the likelihood for chance dispersal e.g., across Makassar Strait being successful by providing a larger 'target' at a lower distance from the source area. A link between Miocene colonization of Sulawesi and sea-level changes was already proposed for squirrels by Mercer & Roth (2003). The effect of sea-level lowstand(s) on decreasing the maximum open water distances between

Sulawesi and the potential source areas would probably have been greatest from the Pliocene onwards (Hall 2009b). The splits of Sulawesi taxa from their sister groups and the earliest speciation events on Sulawesi are particularly densely clustered in the Pliocene/Pleistocene. If this is not merely a sampling artefact, it would support the assumption of an important role for sea-level lowstands coupled with the increasing extent of land area on Sulawesi for dispersal onto the island.

Several taxa have colonized Sulawesi more than once (see Table 2.1), and the time between independent colonization events ranges from almost instantly in geological terms (≤ 0.5 My) to c. 10 My (Fig. 2.5E). The timing of the two colonization events is not correlated to the two geological/climatological events discussed above, i.e., both colonization events in each taxon occurred either in the Miocene or Pliocene. The water beetles (black bars in Fig. 2.5E) are an exception here, but this might be an effect of substitution rate issues (see below) rather than indicative of very differently timed colonizations of Sulawesi. There is also no clear pattern of exclusive distribution of the two Sulawesi lineages in the respective taxa, which might be taken as indicative of the colonization of different potentially separated paleo-islands of Sulawesi. In some taxa different colonizations resulted in a largely or completely overlapping distribution of the Sulawesi clades, e.g., in fanged frogs (Evans et al., 2003a) or shrews (Ruedi et al., 1998), while the two lineages of mite harvestmen for example are found in Southwest and North Sulawesi, respectively (Clouse & Giribet, 2010). However, here as well as in some other cases (Ruedi et al., 1998), the geographic coverage of sampling on Sulawesi is poor. In taxa such as the macaques or the wild boar, the different Sulawesi lineages do not occur sympatrically, but it is doubtful whether monophyly of the Sulawesi taxa can be ruled out, either because there is insufficient support of the respective nodes (wild boars, Larson et al., 2005) or because nuclear data are in conflict with the mtDNA gene trees on this point (macaques, Tosi et al., 2003; Evans et al., 2010). Not only is there no obvious respective pattern in the distribution of lineages that have colonized Sulawesi independently, there is in our opinion also little hard evidence for any Sulawesi clade being confined to part of the island only, which might suggest a link of its origin on Sulawesi and tectonic processes. For most invertebrates, the poor sampling prevents any respective statement, and for vertebrates we are not aware of any recent example (and the fossil record is rather limited; see van den Bergh et al., 2001). Michaux (2010) has proposed that Sulawesi should not be employed as an area of endemism in biographic analyses in its entirety, but that rather parts of different geological or tectonic origin such as West Sulawesi or the Banggai-Sula block should be considered separately. This is not supported by our data. However, subsequent to colonization of the island, geology or climate related changes in paleogeography seem to have played an important role in driving diversification processes on Sulawesi. A pronounced and at least partly congruent geographic structure of Sulawesi populations or species has been found in all taxa which have been sufficiently extensively

sampled on the island (see e.g., Bridle et al., 2001; Evans et al., 2003c, 2008; Larson et al., 2005; Brown et al., 2010; von Rintelen et al., unpubl. data). While not the focus of this study, it would be highly interesting to test for temporal congruence in intra-Sulawesi diversification as well.

The datasets analyzed in this study comprise (mostly non-flying) terrestrial and limnic taxa. The distribution of TMRCA estimates of limnic taxa does not deviate from the general pattern (Fig. 2.5D), suggesting that freshwater organisms are not necessarily suffering "unique biogeographic constraints" (Unmack, 2001). It rather seems that the capability for dispersal varies as much as in terrestrial taxa, and the limnic candidate taxa for a vicariant origin on Sulawesi e.g., comprise taxa absolutely confined to freshwater such as viviparous snails (the pachychilid *Tylomelania*, von Rintelen & Glaubrecht, 2005) and a group of secondary freshwater fishes with brackish water species (telmatherinids, Sparks & Smith, 2004).

Reliability of divergence time estimates

The validity of our results and their interpretation in terms of biogeographic hypotheses is strongly dependent on the quality of the divergence time estimates. This is not an exclusive issue here, as the usage of substitution rates or fossil calibrations for divergence time estimations is crucial in all molecular clock analyses (e.g., Graur & Martin, 2004; Pulquério & Nichols, 2007; Ho et al., 2008; Wilke et al., 2009; Wertheim & Sanderson, 2010). Fossil record data might be imprecise and only give minimum estimates for the age of a group (e.g., Benton & Donoghue, 2007; Donoghue & Benton, 2007); the application of substitution rates derived from non-closely related organisms and/or for different genes is fraught with difficulties (Thomas et al., 2006; Wilke et al., 2009). The latter point is of particular relevance here, since only a subset (n = 10) of the datasets in this study could be analysed using a taxon-specific substitution rate or calibration point, respectively (see Table 2.1 for details and references). The majority of analyses were conducted using a rather unspecific substitution rate (compare Table 2.1), e.g., for insects (general mtDNA rate), mammals (general mtDNA rate), fishes (ND2 rate for cyprinodont fishes) and frogs/toads (ND2 rate for leptodactylid frogs).

Precise calibration points and specific substitution rates are strong priors for the estimation of divergence times and if a large error in either calibration method is unrecognized this makes the resulting divergence time estimates unreliable to an unknown degree (Pulquério & Nichols, 2007). Our study provides striking examples of the issues involved. For the ricefishes, e.g., Takehana et al., 2005 suggested a TMRCA of 29-32 Mya for the Sulawesi lineage compared to the 3-7 Mya estimated here. This discrepancy is largely due to their use of a substitution rate for Antarctic notothenioid fishes (Bargelloni et al., 2000), which is an order of magnitude slower than the rates usually reported for vertebrates (Near, 2004) and does seem an unusual choice for small tropical freshwater fishes. Divergence time estimates also almost

always differed considerably if different genes were analysed for the same taxon (Fig. 2.5F, Table 2.1). The water beetle dataset is rather illustrative here, comprising data from three different gene fragments, which were both independently and jointly (see below) analysed here (datasets 2, 9, 22) using the same rate ('general' insect mtDNA rate of 2.3%/My; Brower, 1994). This approach resulted in extremely varying divergence times for the same Sulawesi taxon (the older of the two lineages found) from its sister group ranging from the Oligocene to the Pleistocene for the individual genes (datasets 2, 9, 22, summarized in Fig. 2.5F). In a recent study by Pons et al. (2010) beetle-specific rates for protein-coding mitochondrial genes have been estimated using a calibration derived from fossils. The rates for some individual genes differ considerably from the 2.3% average of the general insect mtDNA rate. For *cytochrome c oxidase I* (COI, dataset 9, Fig. 2.2), the specific rate is more than eightfold higher than the general insect rate and for *cytochrome b* the rate is more than 30% faster than the general rate (Cyt b, dataset 2, Fig. 2.2). More specific rates thus do not necessarily result in better congruence among the divergence times estimated from different genes (see Table 2.1). Using the new rates, the age of the studied water beetle lineages does decrease considerably for COI, shifting the TMRCA from the Mid-Miocene to the Pleistocene (Fig. 2.2-4, Table 2.1). Pons et al. (2010) also show that the average rates between beetle taxa vary considerably, indicating that a beetle rate may be little better than a general insect rate for estimating divergence times in some lower rank taxa of beetles, which is perhaps not totally surprising in this old and megadiverse 'superradiation' (Hunt et al., 2007). However, other recent studies have demonstrated a good fit between divergence time estimates derived from multiple calibration points and a general rate (Jønsson et al., 2011). In our study, the results from some vertebrate datasets suggest that the use of unspecific rates is not necessarily misleading. For the squirrels (datasets 12, 13), e.g., the calibration point used by Mercer & Roth (2003) was employed and the resulting estimated clock rates (Table 2.2) approximately fit the faster of two published rates for mammals. These were used for all other mammal datasets in this study, such as the tarsiers (dataset 3). Here, the slower mammalian mtDNA rate 1 seems more appropriate, with the mean estimated TMRCA of 8.3 My (see Table 2.2) roughly (considering confidence intervals) fitting the split from *T. bancanus* c. 11 Mya derived from fossil data (Merker et al., 2009), while the faster rate (rate 2 in our study) yields a much younger TMRCA. The substitution rates in primates have long been known to be lower than those of other mammals, particularly rodents (Li et al., 1987) and among primates the lemurs, which are comparatively closely related to tarsiers, have the lowest rates (Hasegawa et al., 1990), so these differences in rate applicability between mammal datasets should perhaps be expected. The application of external, i.e., non-taxon specific, rates in the absence of fossil calibrations is not necessarily misleading, but as it is hard to assess the reliability without external reference, the results must be interpreted with extreme caution.

With the exception of the mite harvestmen, datasets comprising multiple gene fragments were not only analysed separately for each gene but also as a concatenated alignment using one rate (fanged frogs, macaques, rice fishes, water beetles). As expected, the resulting TMRCA was intermediate to the estimates gained from the separate analyses (see Table 2.1). For the macaques and, to a lesser degree, also the water beetles, the TMRCAs from the separate analyses differ strongly and the TMRCA estimated from the concatenated dataset is very far from the mean of the two single gene TMRCAs, even though the same external rate was used in all analyses (Table 2.1). We suspect that this might be an effect of length differences between the individual gene fragments. For the macaques, e.g., the TMRCA of the combined analysis is 18.6 Mya, which is quite close to the 19.5 Mya estimated for the NADH dataset alone, which has a length of 1325 bp and very different from the 3.7 Mya for the 12S dataset with only 413 bp length.

While it has been suggested that a strict clock must be employed in molecular dating analyses using external rates (e.g., Wilke et al., 2009), the latter can also be employed under relaxed clock models as implemented in BEAST (Drummond & Rambaut, 2007). The effect of the choice of molecular clock model (strict versus relaxed clock) on the estimated TMRCAs seems secondary, though, if compared to the differences stemming from applying different substitution rates to the same gene or the same rate for different genes (see results; compare Table 2.1 and Table S2, Supplementary Information). Interestingly, the differences between node age estimates are not linked to the rejection of a strict clock by the Bayes factor analyses in both approaches (compare Table 2.2 and Table S2, Supplementary Information); for the tarsiers, e.g., a strict clock was not rejected by the Bayes factor values of the DHM analysis nor was the Bayes factor (BEAST) remarkably high (Table 2.2). The choice of clock model clearly matters and a strict clock model should not be used when rejected by Bayes factors, but we suggest that the choice of substitution rate and calibration is generally more important with respect to the robustness of divergence time estimates. Remarkably, the biogeographic implications of our analyses remain valid under each model or rate used.

Biogeographic interpretation of molecular divergence time estimates

Our study essentially relies on comparing divergence times of Sulawesi taxa with vicariant events derived from geology (Fig. 2.1). This approach has led to a 'resurrection of oceanic dispersal' (de Queiroz, 2005; Cowie & Holland, 2006; Heaney, 2007) in historical biogeography by using divergence time estimates to falsify vicariance hypotheses by demonstrating that the phylogenetic splits in question postdate a vicariance event (e.g., Knapp et al., 2005; Heinicke et al., 2007; Michalak et al., 2010; Schweizer et al., 2010; Bartish et al., 2011). Crisp et al. (2011) have recently argued to use calibrated trees in biogeography in a more explicitly hypothesis-driven way instead of an inductive 'pattern first, process later' procedure

of fitting some geological data to the tree. Overall, our study provides ample evidence in justification of their argument. Crisp et al. (2011: Box 1) have also suggested that a vicariance scenario is only supported by the data if the timing of respective phylogenetic splits and their confidence is overlapping with the time frame of the vicariant events. While we acknowledge that this approach is rigorous and objective, we would suggest that a TMRCA predating a vicariant event, as in the case of the mite harvestmen (Fig. 2.2, dataset 1) in this study, does not automatically falsify an involvement of vicariance in causing the phylogenetic split. The mite harvestmen are probably representative for many groups of small tropical invertebrates in having a high likelihood of being both taxonomically and geographically undersampled (compare Clouse & Giribet, 2010: Fig. 2.1). Extinction could also cause a time lag between the split recovered in a phylogeny of extant taxa and the vicariant event. It might be argued that the risk of extinction is particularly high in Wallacea and adjacent regions where on-going tectonic processes entail frequent local submergence and uplift events.

While stem age estimates are widely used to test vicariance hypotheses in biogeography (see Crisp et al., 2011 and above), this approach has recently been heavily criticized (Heads, 2005, 2009, 2011; Nelson & Ladiges, 2009), particularly the use of divergence time estimates as maximum ages for phylogenetic splits. If a molecular clock has been calibrated using fossils or external rates ultimately derived from fossil calibrations, any resulting age estimate will be a minimum age (e.g., Donoghue & Benton, 2007). Using this age estimate as a *de facto* maximum age which is then employed in testing congruence between the age of a vicariant event and a specific phylogenetic split can indeed create problems if the TMRCA of a lineage is much older than the fossil used to define its age or if the difference between the age geological vicariant event and the phylogenetic split is small (though confidence intervals in a properly conducted molecular clock analysis would then probably overlap with the time range assigned to the vicariant event). Heads (2009, 2011) suggested relying on tectonic calibrations and use well established vicariance events instead of fossils for fixing node age. We consider this proposal even more problematic than the issue it tries to overcome, mainly for two reasons: (1) age estimates for vicariant events can be imprecise as well, as also acknowledged by Heads (2011) e.g., for the Panama Isthmus, and, more importantly, (2) using them would either lead to circular reasoning if the aim of a study is to test for a role of vicariant events in the evolution of a taxon (see also (Kodandaramaiah, 2011) or imply the *a priori* acceptance of a vicariance model, which would make biogeography less of a science and more of a religion. While naive dispersal concepts have been rightly criticized (e.g., Croizat et al., 1974), which paved the way for vicariance biogeography (e.g., Wiley, 1988), it seems futile to deny a role of either dispersal (long-distance dispersal) or vicariance in biogeography (for details see (Thornton, 1983).

Conclusions

Sulawesi has a fauna of mixed geographic and temporal origin. Given the patterns discussed here, it is little surprising that Wallace had difficulties in understanding the biogeography of this "anomalous island" (Wallace, 1876). Dispersal seems the primary mechanism of bringing taxa to the island and a standardized molecular clock approach has led to the falsification of vicariance hypotheses for some Sulawesi taxa of Asian origin. The divergence time estimates presented here are compatible with a vicariance hypothesis in 10-20% of the studied taxa, though, depending on the choice of substitution rate in two cases. A vicariance scenario is dependent on accurate tectonic and paleogeographical data, but evidence for vicariance can at the same time inform geology and e.g., increase the probability for assuming land in certain parts of Wallacea or suggest where geological evidence for this might be sought. Sulawesi has a rich fauna (and flora), and we have used but a few taxa for which the respective data were available for this review of the biogeography of the island. Despite the methodological problems inherent to a molecular clock approach and the by necessity poor taxon coverage we believe our results are generally robust. However, more data are needed to gain a truly comprehensive understanding of Sulawesi's biogeography, including intra-island diversification. Biogeographers in the pre-molecular and pre-cladistic age compiled many fascinating distribution patterns involving Sulawesi, and the simplified 'origin from Asia' or 'origin from Australia' dichotomy this study is largely confined to will surely become more complex. Recently, e.g., the dating of phylogenetic splits in two genera of nymphalid butterflies has revealed sister groups of the Sulawesi lineages from Timor and the Moluccas, respectively (Müller & Beheregaray, 2010; Müller et al., 2010). The time frame of Sulawesi colonization for the butterflies (Miocene) and speciation on the island (Miocene?/Pliocene) is fully compatible with the data presented here, though, and suggests that Sulawesi's fauna was decisively shaped by events in the Miocene and Pliocene. The need for new data also applies to geology. Biogeographers must acknowledge the limitations of tectonic or paleogeographic reconstructions (Hall, 2009a), but they should also use them explicitly to derive temporal constraints for biogeographic hypotheses of vicariance, as e.g., exemplarily done for the Madagascan fauna by Warren et al. (2010). This will only work in conjunction with molecular divergence time estimates. Again, these are not the Holy Grail but biological data with a confidence margin and should be treated accordingly, but ignoring them with a vague reference to paleoendemics (Michaux, 2010) is not an option in our opinion.

Chapter 3

Age estimates for an adaptive lake fish radiation, its mitochondrial introgression, and an unexpected sister group: sailfin silversides of the Malili Lakes system in Sulawesi

Introduction

Adaptive radiations of organisms restricted to habitat islands are among the prime model systems for investigating speciation processes in nature (Schluter, 2000; Losos & Mahler, 2010; Santos & Salzburger, 2012; Martin & Wainwright, 2013). Estimations of rates of radiation, as well as the reconstruction of past geographic scenarios of divergence, require the incorporation of a temporal axis into phylogenetic reconstructions (e.g., Lerner et al., 2011; Friedman et al., 2013; Bess et al., 2014). Likewise, the dating of phylogenetic splits may contribute to the understanding of other processes potentially contributing to evolutionary divergence (and adaptive radiation), such as introgressive hybridization (Schwarzer et al., 2011; Willis et al., 2014).

The ancient lakes of Sulawesi's central highlands are a hotspot of aquatic diversity, strongly dominated by endemic species flocks, including radiations of freshwater fishes (von Rintelen et al., 2012). The Malili Lakes system (Matano, Mahalona, and Towuti) – a hydrological chain of three main lakes interconnected by rivers – is the main habitat of the sailfin silversides radiation (Atherinomorpha: Telmatherinidae). About 30 morphospecies of these sexually dimorphic fishes that typically possess bright male colourations are distinguished in the Malili Lakes drainage and a few surrounding rivers (Herder et al., 2006b). Sailfin silversides show conspicuous colour polymorphisms (e.g., Gray et al., 2008; Pfaender et al., 2013), and there are clear indications that ecological speciation has shaped their adaptive radiation (e.g., Pfaender et al., 2010, 2011; Cerwenka et al., 2012). Phylogenetic analyses suggest that the two lineages of *Telmatherina* radiating in the hydrological head of the lakes system, ancient graben-lake Matano, form an ancient monophyletic group that was later compromised by introgressive hybridization from stream populations (Herder et al., 2006a). This introgression has affected only the 'sharpfins', a lineage of predominantly epibenthic sailfin silversides, whereas their rather pelagic sister group, the 'roundfins', show no indications of genetic exchange with stream populations (Herder et al., 2006a; Schwarzer et al., 2008; see Herder & Schliewen, 2010 for discussion).

Morphological data support a clade composed of *Marosatherina ladigesi* from south-west Sulawesi, and *Kalyptatherina helodes* from the islands Batanta and Misool off the Vogelkop peninsula (Birds' Head, New Guinea), as sister group of the Telmatherinidae in and close by the Malili Lakes (Aarn et al., 1998). However, only *Marosatherina* has been considered as an outgroup to the lakes radiation in genetic studies so far (Herder et al., 2006b; Roy et al., 2007). Likewise, Telmatherinidae are represented only by *Marosatherina*, a species available worldwide by the aquarium trade, in most phylogenetic studies targeting relationships within the Atherinomorpha (e.g., Sparks & Smith, 2004; Setiamarga et al., 2008; Unmack et al., 2013).

Here, we use mitochondrial data to (i) assess the relationship of *Kalyptatherina* relative to *Marosatherina* and the Malili Lakes Telmatherinidae, (ii) estimate the age of the sailfin silverside radiation of the Malili Lakes, and (iii) provide an estimation of the age of the mitochondrial introgression present in Lake Matano's sharpfin *Telmatherina*. For this, we combine sequence data of the Telmatherinidae with data of representatives of the Melanotaeniidae, the closely related rainbowfishes from Australia and New Guinea (Parenti, 1993; Sparks & Smith, 2004; Setiamarga et al., 2008; Unmack et al., 2013), and estimate divergence times using both indirect and geological calibration points, and substitution rates suggested by Unmack et al. (2013) as telmatherinid fossils are missing.

Materials and Methods

DNA extraction, amplification and sequencing

DNA from 99 specimens, representing 74 taxa or populations (Table S1), was extracted using the QIAGEN DNeasy® Blood & Tissue Kit following the manufacturers' instructions. Two mitochondrial loci were amplified by using the Sigma Taq-Polymerase system: partial NADH dehydrogenase subunit 2 (ND2, 830 bp length) and a combined fragment comprising partial 12S rRNA, tRNA-Val, and partial 16S rRNA ('12S-16S fragment'; ca. 1,275 bp length). All PCR reactions were conducted using the same conditions: 6 min at 94 °C (initial denaturation); 45 cycles: 1 min at 94 °C (denaturation); 1 min at 45 °C (annealing); 1.5 min at 72 °C (elongation). Two sets of primer pairs (ND2) and three different primers (12S-16S fragment) were used for both amplification and sequencing (Table 3.1). The 12S-16S fragment could not be amplified for all *Kalyptatherina helodes* specimens; the same applies to the ND2 fragment for *Melanotaenia lacustris* (AP00419) and *Melanotaenia boesemani*. Missing data was replaced by Ns for these specimens. PCR products were purified using the enzymatic digestion system USB ExoSAP-It. Sequencing was carried out by the Sequencing Service of the Ludwig-Maximilians-Universität Munich, Department of Genetics using a ABI 3730 48 capillary sequencer. All sequences are deposited in GenBank at the NCBI [ND2, GenBank: KJ667866-KJ667963; 12S-16S, GenBank: KJ667771-KJ667865].

Table 3.1. Primers used in the present study (see Methods for PCR conditions).

Primer name	Sequence	Source
ND2Met	5'-CAT ACC CCA AAC ATG TTG GT-3'	Kocher et al. (1995)
ND2Trp	5'-GTS GST TTT CAC TCC CGC TTA-3'	Kocher et al. (1995)
ND2Gln	5'-CTA CCT GAA GAG ATC AAA AC-3'	Kocher et al. (1995)
ND2Asn	5'-CGC GTT TAG CTG TTA ACT AA-3'	Kocher et al. (1995)
12SF1	5'-TGA AGG AGG ATT TAG CAG TAA G-3'	Rüber et al. (2003)
12SF2	5'-TCT CTG TGG CAA AAG AGT-3'	Rüber et al. (2003)
16SR1	5'-AAG TGA TTG CGC TAC CTT CGC AC-3'	Rüber et al. (2003)

Phylogenetic analyses and estimation of divergence times

Single gene sequences were aligned using MAFFT (Katoh & Toh, 2008; default settings; http://www.ebi.ac.uk/Tools/msa/mafft) and corrected by eye; ambiguous sites at the 5' end of the 12S-16S fragment were removed manually. Both single gene alignments were concatenated using SequenceMatrix v. 1.7.8. (Vaidya et al., 2011) resulting in a total alignment of 2,102 bp. The GTR+G substitution model was used for maximum likelihood (ML) analyses using RAxML BlackBox (Stamatakis et al., 2008; partition model; 100 bootstrap replicates) and Bayesian inference (BI) using MrBayes v. 3.1.2 (Ronquist & Huelsenbeck, 2003; partition model; ngen = 10,000,000; nchains = 4; samplefreq = 500; burnin = 10,001). *Iso rhotophilus* was used as outgroup in all analyses as suggested in Bloom et al. (2012); phylograms were visualized and re-rooted in FigTree v. 1.4 (available at: http://tree.bio.ed.ac.uk/software/figtree).

Single-gene datasets were reduced to unique haplotypes using DAMBE v. 5.1.1 (Katoh & Toh, 2008). Substitution models were selected using jModelTest (Posada, 2008; AIC and BIC selected GTR+G for both partitions). Divergence times were estimated in BEAST v. 1.7.4 (Drummond et al., 2012) on the CIPRES Science Gateway web portal (Miller et al., 2010). Each analysis was run twice using the following settings: separate partitions used; ingroup monophyly enforced; uncorrelated lognormal relaxed-clock model; ngen = 40,000,000; samplefreq = 2,000; burnin = 10,001; Yule model; calibration points with normal distribution; the RAxML tolopolgy was used as starting tree for all analyses. Four different analyses were: [A] indirect calibration of the root height using the estimated divergence time of the split of *Iso–Melanotaenia* obtained from Setiamarga et al. (2009) (93-113 Mya): Root height was set to 106.0 +/- 10.0 My; [B] indirect calibration of the melanotaeniid northern–southern split inferred from a recent molecular clock analysis by Unmack et al. (2013) (23.8-30.8 My): Node 9 (see Fig. 3.1), representing the split of northern and southern clades of New Guinean melanotaeniids was set to 27.3 +/- 3.5 My; [C] a separate run was performed using a substitution rate of 1%/My for mtDNA suggested and used by Unmack et al. (2013); [D] geological calibration of the Central Highlands formation in New Guinea (10-14 Mya; Hill & Hall, 2003) putatively resulting in the geographical separation of the two major melanotaeniid clades ('northern' and 'southern'; see also Unmack et al., 2013): Node 9 (see Fig. 3.1), representing the split of northern and southern clades of New Guinean melanotaeniids and is probably related to the abovementioned orogeny was set to 12 +/- 2.0 My. See Table 3.2 for results among datasets. Log files of replicates were visualized in Tracer v. 1.5 (Rambaut & Drummond, 2007) for congruency and combined in LogCombiner v. 1.7.4 (BEAST package; 50% burnin). Log files were visualized in Tracer v. 1.5; ESS values for each parameter never dropped below 200 except for analysis [C] (low values for some parameters such as 'prior' and 'posterior', however 'likelihood' is >1,000). Tree files were combined in LogCombiner v. 1.7.4 (50% burnin) and summarized in TreeAnnotator v. 1.7.4 (BEAST package; no burnin; MCC trees = maximum clade credibility trees).

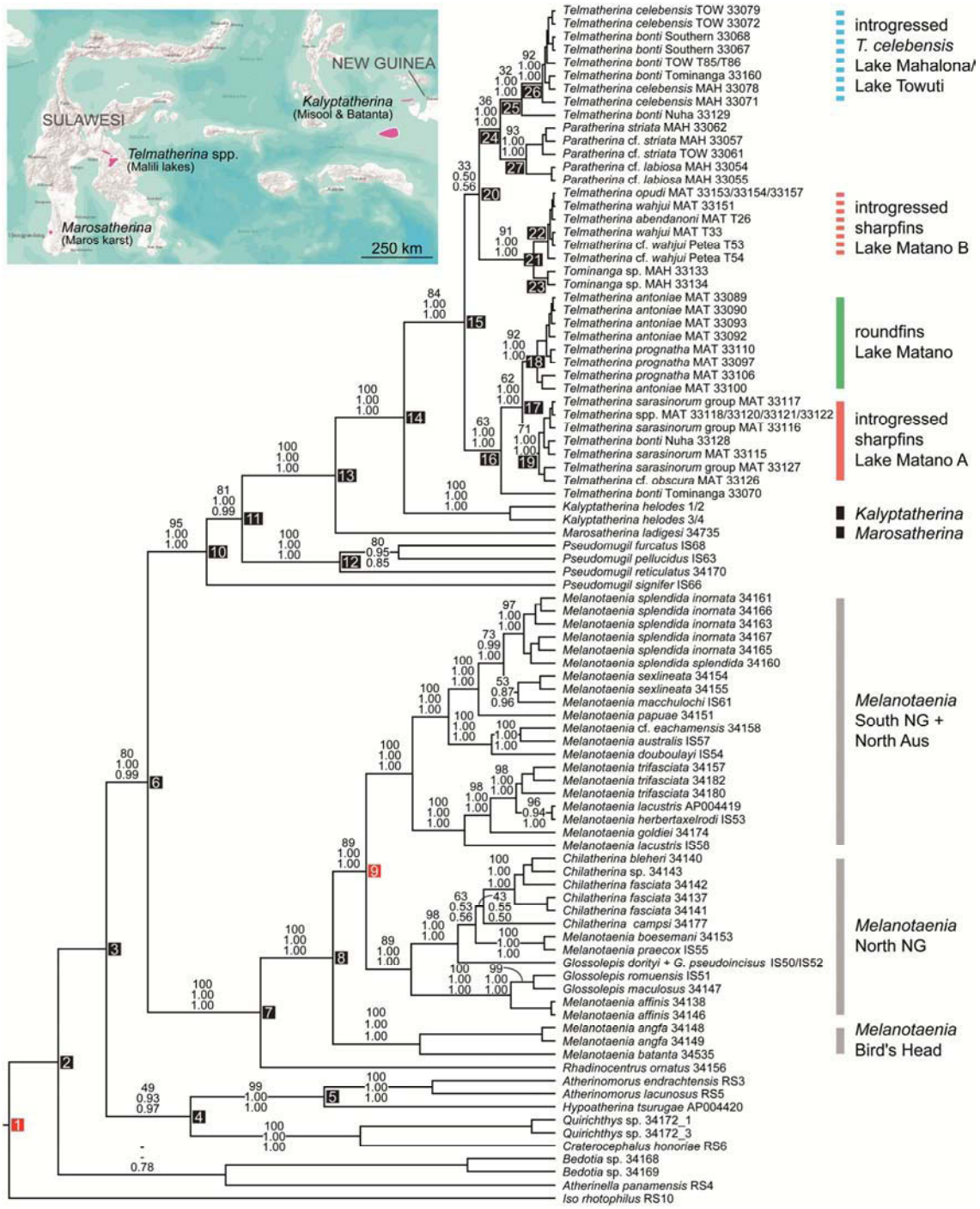


Figure 3.1. BEAST MCC tree showing general melanotaeniid and telmatherinid relationships for the concatenated dataset. Topology derived from analysis [D]. Numbers on branches denote RAxML bootstrap values, MrBayes posterior probabilities and BEAST posterior probabilities (from top to bottom); numbers on nodes correspond to node numbers given in Table 3.2, i.e., divergence time estimates across all four analyses performed. Nodes 1 and 9 were used as indirect and geological calibration points, respectively (see Methods for details). Abbreviations used for sailfin silversides: MAH = Lake Mahalona; MAT = Lake Matano; TOW = Lake Towuti. See text and Fig. 3.3 for details on the Lake Matano telmatherinid radiation. Map shows current distribution of *Kalyptatherina*, *Marosatherina*, and the Malili Lakes telmatherinds.

Table 3.2. Divergence time estimates for selected nodes.

	Mean age (lower and upper 95% HPD) [My]			
Analysis	[A]	[B]	[C]	[D]
Node				
1	103.09 (82.70, 122.15)	70.36 (39.70, 106.48)	33.15 (23.43, 43.20)	31.31 (16.40, 48.90)
2	93.46 (69.89, 116.89)	64.19 (36.83, 95.98)	30.46 (22.53, 39.22)	28.65 (14.92, 43.70)
3	83.62 (59.92, 107.04)	57.69 (33.54, 85.19)	27.47 (20.40, 34.95)	26.03 (13.56, 39.08)
4	67.20 (40.38, 93.53)	46.10 (22.62, 70.57)	21.88 (13.85, 31.38)	20.40 (9.07, 32.21)
5	41.45 (18.04, 65.92)	22.64 (8.27, 39.89)	10.44 (3.89, 17.11)	9.95 (3.14, 18.18)
6	75.85 (53.25, 98.86)	52.00 (29.69, 76.73)	24.94 (18.31, 31.63)	23.56 (12.17, 35.75)
7	54.59 (35.08, 75.21)	39.49 (24.69, 56.03)	18.74 (13.30, 24.86)	17.41 (9.49, 26.50)
8	40.78 (24.86, 58.33)	30.34 (20.55, 40.64)	13.09 (9.22, 13.73) | 32.7 (28.4, 37.4)	12.74 (7.54, 18.43)
9	34.50 (20.65, 49.72)	25.85 (18.83, 32.88)	11.04 (7.69, 14.78) | 27.0 (23.8, 30.8)	10.87 (6.70, 14.98)
10	64.76 (43.47, 86.22)	43.85 (25.27, 65.90)	21.35 (15.20, 27.93)	20.19 (10.38, 30.97)
11	57.80 (37.80, 78.33)	38.95 (22.54, 59.33)	19.10 (13.38, 25.44)	18.11 (9.05, 27.98)
12	37.88 (18.74, 56.83)	25.00 (11.00, 39.86)	13.15 (6.95, 19.24)	12.30 (4.90, 20.65)
13	42.86 (25.60, 62.34)	27.17 (14.37, 42.94)	13.60 (8.89, 18.66)	12.93 (6.18, 20.36)
14	28.90 (15.01, 43.48)	19.15 (9.56, 30.43)	9.11 (5.50, 13.06)	8.45 (3.79, 13.45)
15	17.43 (8.21, 27.65)	11.99 (5.78, 19.06)	5.35 (3.22, 8.12)	5.18 (2.15, 8.53)
16	9.59 (3.52, 17.08)	7.13 (3.21, 12.20)	3.12 (1.47, 4.91)	3.00 (1.01, 5.28)
17	6.43 (2.40, 11.60)	4.45 (1.95, 7.62)	1.98 (0.90, 3.20)	1.86 (0.63, 3.27)
18	3.66 (1.18, 7.03)	2.45 (0.81, 4.31)	1.04 (0.39, 1.82)	1.01 (0.31, 1.92)
19	3.05 (0.90, 5.81)	2.24 (0.73, 4.15)	0.99 (0.36, 1.78)	0.93 (0.27, 1.80)
20	13.94 (6.35, 22.44)	10.15 (4.84, 16.54)	4.40 (2.32, 6.70)	4.32 (1.72, 7.17)
21	3.60 (0.73, 7.77)	2.92 (0.64, 6.29)	1.27 (0.33, 2.68)	1.22 (0.25, 2.68)
22	1.21 (0.22, 2.61)	0.92 (0.17, 1.98)	0.40 (0.09, 0.83)	0.37 (0.07, 0.83)
23	1.20 (0.08, 2.97)	0.96 (0.08, 2.44)	0.41 (0.04, 1.00)	0.38 (0.02, 0.95)
24	10.23 (4.58, 17.24)	7.24 (3.09, 11.98)	3.18 (1.54, 5.07)	3.03 (1.19, 5.33)
25	6.14 (2.08, 11.43)	4.39 (1.44, 8.03)	1.92 (0.73, 3.42)	1.82 (0.48, 3.45)
26	2.36 (0.64, 4.63)	1.68 (0.50, 3.26)	0.73 (0.23, 1.34)	0.70 (0.18, 1.37)
27	5.05 (1.54, 9.36)	3.84 (1.09, 7.02)	1.67 (0.61, 3.02)	1.61 (0.48, 3.07)
Resulting rate [% / My]	0.33 (ucld.mean)	0.47 (ucld mean)	1.00 (ucld mean)	1.14 (ucld mean)

Grey-marked cells denote nodes used for prior calibration; [A]: root height = 106.0 +/- 10.0 My; [B]: 27.3 +/- 3.5 My; [C]: relaxed clock = 1%/My – node 8 and 9 show striking differences between the present analysis (first column) and the node ages inferred by Unmack et al. (2013) (second column); [D]: 12.0 +/- 2.0 My. See Methods and Fig. 3.1 for details.

Results

Phylogenetic relationships and age estimates

The phylogenetic reconstructions show a well to highly supported sister group relationship between Melanotaeniidae and a clade comprising species of the genus *Pseudomugil* (Pseudomugilidae) and all representatives of the family Telmatherinidae (Fig. 3.2 and node 6 in Fig. 3.1). Inferred mean ages for the split (node 6) range from 23.6-75.9 My among the four molecular analyses performed in BEAST (Table 3.2). Basal relationships among atherinid, atherinopsid and bedotiid species are not well resolved in ML, BI and BEAST analyses, and estimated mean ages for these basal nodes are strikingly different and range between 31.3 and 103.1 My for the root height (node 1) and the first basal split (node 2: 28.7-93.5 My).

The Melanotaeniidae form a highly supported monophyletic group, including *Rhadinocentrus ornatus* from Queensland, and all species of the genera *Melanotaenia*,

Chilatherina and *Glossolepis* analyzed. In this group, *R. ornatus* represents the sister taxon to three distinct and highly supported clades from New Guinea (and surrounding islands), and northern Australia (Fig. 3.1). Interestingly, *Glossolepis dorityi* and *G. pseudoincisus* share one haplotype (Fig. 3.1), which is most likely due to gene flow between the two species inhabiting the very same river system in the northern lowlands of New Guinea. Two *Melanotaenia* species from the Bird's Head (*M. angfa*) and Batanta Island (*M. batanta*) form a monophyletic group, being sister to two separate clades comprising *Melanotaenia* (sub)species from southern New Guinea and northern Australia, and species of the genera *Chilatherina*, *Glossolepis* and *Melanotaenia* from northern New Guinea (cf. 'western', 'southern', and 'northern' clade in Unmack et al., 2013; Fig. 3.1). This is largely congruent with findings of Unmack et al. (2013). The inferred mean age estimates for the western–northern/southern split (node 8 in Fig. 3.1) and the northern–southern split (node 9 in Fig. 3.1) range from 12.7-40.8 and 10.9-34.5 My, respectively, depending on if node 9 was enforced in the respective analysis (see Table 3.2, analyses [B and D]).

Phylogenetic reconstructions (ML and BI) show a well to highly supported clade comprising members of the genus *Pseudomugil* (*P. signifer*, *P. reticulatus*, *P. furcatus*, and *P. pellucidus*; note that the Pseudomugilidae are paraphyletic; see also below), and the Telmatherindae. Within the Telmatherinidae, *Marosatherina ladigesi* from the Maros karst (southwest Sulawesi) is basal to the clade including *Kalyptatherina helodes* (Batanta and Misool) plus the sailfin silversides from the Malili Lakes system in Central Sulawesi, Indonesia (node 14; mean age range: 8.5-28.9 My). Estimated divergence times range from 18.1-57.8 My for the split of Telmatherinidae and the most recent Pseudomugilidae clade (node 11), and are quite similar for the first diversification events within each family (12.3-37.9 My for node 12 for the most recent *Pseudomugil* clade, and 12.9-42.9 My for node 13, respectively).

In contrast, species of the genus *Telmatherina* fall in three distinct mitochondrial clades. Node 25 (1.8-6.1 My) represents the MRCA of a clade (from now on we refer to MRCAs of particular clades when discussing nodes) is composed of the lake-dwelling *Telmatherina celebensis* from Lakes Mahalona and Towuti, and several populations of stream-dwelling *Telmatherina bonti*. Node 22 (0.4-1.2 My) comprises sharpfin specimens from Lake Matano, together with stream-dwellers; node 16 contains the remaining *Telmatherina* from Lake Matano, namely sharpfins (node 19), roundfins (node 18), and two stream *Telmatherina* from River Nuha (north of Matano) and River Tominanga (draining Lake Mahalona to Towuti). Taken together, these results are consistent with those reported by Herder et al. (2006a). In the light of nuclear and morphological data (cf. Herder et al., 2006a), the mitochondrial signatures provide evidence for substantial hybridization between lake- and stream-dwelling *Telmatherina*. This results in two haplotype clades, the 'original' Matano haplotypes (node 19) being sister to

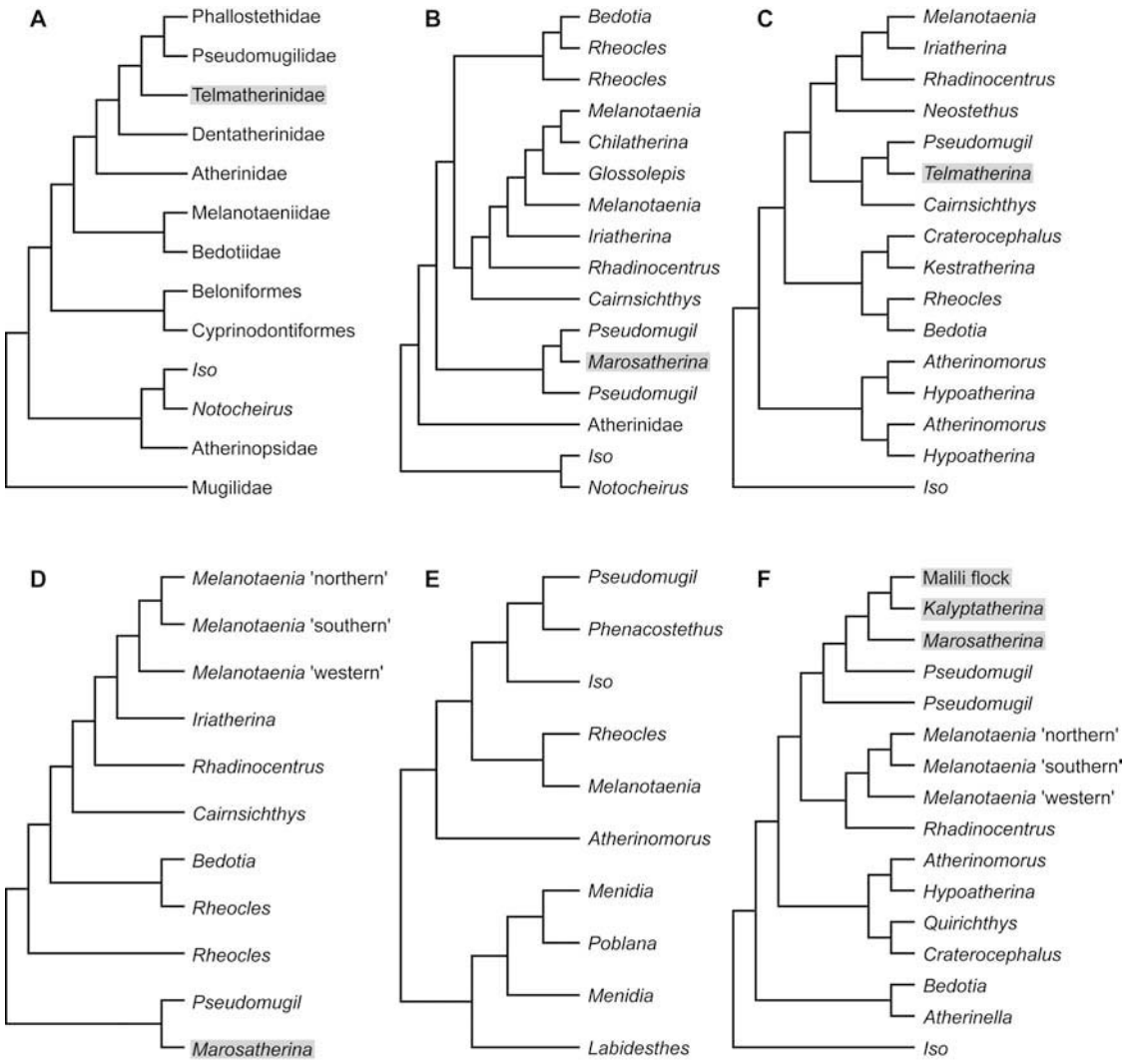


Figure 3.2. Phylogenetic hypotheses showing the systematic position of Telmatherinidae and relatives among different studies. (A) Saeed et al. (1994), (B) Sparks & Smith (2004), (C) Bloom et al. (2012), (D) Unmack et al. (2013), (E) Near et al. (2013), (F) Present study. Telmatherinid representatives are highlighted in grey. See text for details.

the endemic roundfins (node 18), and the introduced haplotypes of (node 22) within Lake Matano's sharpfins.

The comparison of diversification and mitochondrial introgression events within the Lake Matano radiation suggests that the ancient clade of haplotypes, endemic to the lake (node 17), is approximately 1.9-6.4 My old. It comprises the clades of the lake's two sub-radiations, namely the pelagic to benthopelagic roundfins (node 18: 1.0-3.7 My), and the 'native' sharpfin haplotypes of the lake (node 19: 0.9-3.1 My). In contrast, the age of the introgressed haplotypes present in parts of the sharpfin individuals (node 22) is apparently more recent (0.4-1.2 My). A comparatively recent origin of mitochondrial introgression is also supported for the second case of massive hybridization of stream- and lake-haplotypes: node 26, representing the *Telmatherina celebensis* clade from Lakes Mahalona and Towuti, is estimated to only 0.7-2.4 My (see Fig. 3.3 and Table 3.2 for details).

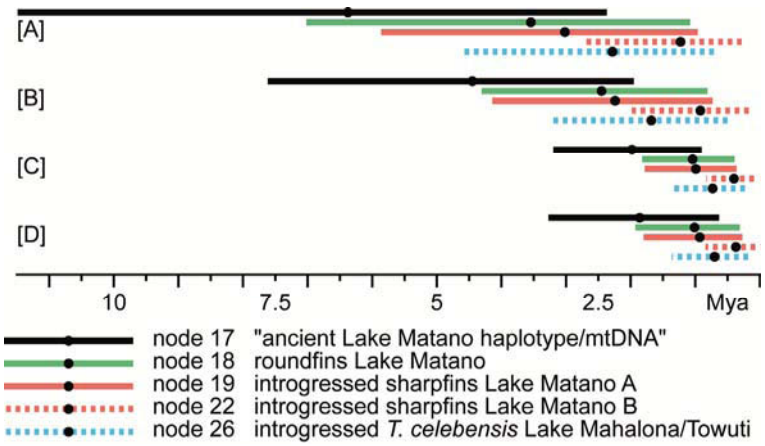


Figure 3.3. Relaxed-clock divergence times distribution among analyses for the Lake Matano telmatherinid radiation. Bars denote 95% credibility intervals, dots represent mean ages for MRCAs of the respective clades. See Methods and Table 3.2 for details.

Discussion

The endemic sailfin silversides radiation of the Malili Lakes serves as a model system in evolutionary ecology research (see Herder & Schliewen, 2010; Vaillant et al., 2011; von Rintelen et al., 2012 for reviews). The detailed and temporal reconstruction of the phylogenetic history of the *Telmatherinidae* is crucial for exploring the likely conditions underlying speciation processes, including the spatial origin and morphological traits of ancestral and introgressed populations. It remained, however, unclear if it is justified to consider *Marosatherina* from SW Sulawesi as the sister species to the lakes radiation, and if the age of the lakes radiation in fact falls into preliminary lake age estimations proposed by geologists and limnologists, which are however not yet fully reliable (Vaillant et al., 2011). These preliminary seismic data suggest an age of at least 600,000-700,000 years for Lake Towuti (Vaillant et al., 2011), while the geological fault formation, in which Lake Matano is embedded, might be comparatively older possibly starting around 4 Mya (Wilson & Moss, 1999). An onset of the Malili Lakes system formation in the early Pleistocene c. 1-2 Mya is plausible (Robert Hall, pers. comm.).

The closest relative to the sailfin silversides radiation

The phylogenetic reconstructions presented here are based on mitochondrial markers only, and hence do not allow inferences about cyto-nuclear discordances within the sailfin silversides species flock (cf. Herder et al., 2006a). However, this mtDNA dataset allows inferring a first hypothesis for reconstructing phylogenetic relationships among disjunctly distributed

telmatherinid species, whose mtDNA haplotypes most likely have preserved the vicariant phylogenetic signal.

Our analyses clearly suggest that mtDNA haplotypes of the Malili Lakes radiation are more closely related to *Kalyptatherina helodes,* the only telmatherinid species occurring on islands east of Sulawesi (Misool and Batanta), rather than to *Marosatherina ladigesi* from SW Sulawesi, a species previously considered the sister taxon to the Malili flock. Its inclusion into the Telmatherinidae is highly plausible in the light of morphological data (Saeed & Ivantsoff, 1991), but its placement as the sister taxon to the lacustrine flocks appears surprising. However, reconstructions of the complex geological history of Sulawesi and adjacent islands may provide explanations for these findings (see section below).

The phylogenetic relationships of the families within Atheriniformes, and the composition of these families, remain partially controversial, most likely due to substantial differences in taxon sampling and the methods applied (Fig. 3.2). Two important taxa could not be included in the present dataset, namely *Cairnsichthys rhombosomoides* (Queensland, Australia), and the rather widespread brackish water family Phallostethidae (priapumfishes; recorded from Sundaland, Luzon, Palawan and Southwest Sulawesi; Parenti, 1996; Parenti & Louie, 1998). *Cairnsichthys* is suggested to be basal to *Pseudomugil* and *Telmatherina* according to the molecular phylogeny by Unmack et al. (2013), while the phallostethids are placed as the sister to *Pseudomugil* according to Near et al. (2013). Morphological studies by Saeed et al. (1994) suggested Melanotaeniidae as sister to Madagascar's Bedotiidae, which are in turn most closely related to a clade composed of Telmatherinidae, Pseudomugilidae (blue eyes), and three other atheriniform families. The combined analysis of molecular and morphological data by Sparks & Smith (2004) supported the sister group relationship of monophyletic Melanotaeniidae and Bedotiidae; a clade composed of *Marosatherina* (Telmatherinidae) and two *Pseudomugil* species (Pseudomugilidae) represent the sister clade to Melanotaeniidae and Bedotiidae in that study. A comprehensive study by Bloom et al. (2012) reported contrasting results, based on the combined analysis of one mitochondrial and nuclear marker. The single representatives of *Pseudomugil* and *Telmatherina* analyzed by Bloom et al. (2012) were also supported as most closely related, but nested within a clade composed of Melanotaeniidae and Phallostethidae. Most recently, Near et al. (2013) confirmed the inclusion of Phallostethidae within the atherinomorphs as distant sister to a monophyletic group of *Pseudomugil* species, based on 10 nuclear markers. Unmack et al. (2013) found in a comprehensive multilocus molecular study (seven mtDNA markers and one nuclear marker) that Melanotaeniidae are the monophyletic sister group to Madagascar's *Rheocles* and *Bedotia*; Melanotaeniidae were again supported as sister to a clade composed of *Marosatherina* and *Pseudomugil*. The results of the present study clearly support the close relationship between Telmatherinidae and Pseudomugilidae reported by previous workers. However, the

Pseudomugilidae are also clearly paraphyletic, with *P. signifer* being basal to all remaining Pseudomugilidae and Telmatherinidae (Fig. 3.1). This finding is consistent with the results of Sparks & Smith (2004) (see also Fig. 3.2), and highlights the need for a detailed study addressing the systematic position of several *Pseudomugil* species, especially that of *P. signifer*.

Biogeographic implications

Divergence time estimates of the split of *Kalyptatherina* and the central Sulawesi Malili flock of 8.5-28.9 Mya renders a scenario of terrane-rafting the most plausible explanation for the present-day distribution pattern, given the region's geological history. In contrast, and considering its present local distribution and estimated node ages, the ancestor of *Marosatherina* most likely colonized western Sulawesi by dispersal from the Sahul Shelf, the current distribution of *Pseudomugil* and Melanotaeniidae.

West Sulawesi (i.e., the West Sulawesi Plutono-Volcano Arc), was separated from the Asian margin when the Makassar Strait opened in the Eocene c. 45 Mya (Moss & Wilson, 1998; Hall, 2009b; Spakman & Hall, 2010). Extension and westward movements of the Sula Spur (a large promontory of the Australian margin) resulted in a collision with the North Sulawesi volcanic arc c. 20-23 Mya, and a final amalgamation of the Sula Spur (comprising East Sulawesi, Central Sulawesi and Banggai-Sula) with West Sulawesi in the Miocene c. 10-20 Mya (Fig. 3.4; see Lohman et al., 2011; Stelbrink et al., 2012 for more details on general terrane movements and distribution of land and sea in the Indo-Australian Archipelago, and particularly in Sulawesi). Geological separation and fusion represent constraints on vicariant processes involving the colonization of the island's terrestrial and freshwater biota.

A scenario of 'terrane-rafting' provides a plausible explanation for the sister group relationship between *Kalyptatherina*, endemic to the small islands off the Vogelkop Peninsula of New Guinea, and the Malili Lakes sailfin silversides. Geological elements formerly belonging to the Sula Spur were in proximity to the Australian margin – including old offshore islands like Batanta and Misool – before this promontory was extended, moved westwards, and finally collided with West Sulawesi (e.g., Hill & Hall, 2003; Hall, 2009b). It appears most plausible that the population ancestral to the Malili sailfin silversides originates from the Sahul Shelf area, and was dispersed on such a 'terrane raft' when the Sula Spur was extended and moved westwards until this fragment ('raft') collided with West Sulawesi. However, given the temporal uncertainties in both the separation of the Sula Spur from the Sahul Shelf (c. 15 Mya; see also Stelbrink et al., 2012) and divergence time estimates among the four analyses, it remains difficult to test this hypothesis. Mean ages and credibility intervals suggest that this scenario might be plausible for analyses [A] and [B] (15.0-43.5 My and 9.6-30.4 My; see Table 3.3), while the credibility intervals obtained from analyses [C] and [D] (5.5-13.1 and 3.8-13.5) would slightly post-date the estimated age of the Sula Spur separation. However, the lack of

fossil remains requires denser sampling and the incorporation of multiple markers to explain the relationships of present-day geographically distant groups, which probably have been in vicinity in the past.

A marine dispersal explanation for the sister group relationship between *Marosatherina*, a monotypic genus endemic to the Maros karst in SW Sulawesi, and all the remaining sailfin silversides, appears most likely based on its current distribution and the divergence time estimates inferred (node 13: 12.9-42.9 My). West Sulawesi and the remaining geological parts of the island are of different origin (Sunda Shelf and Sahul Shelf), and amalgamated not until in the Miocene due to tectonic movements. It appears plausible to assume that the population ancestral to *Marosatherina* might have originated in the Sahul Shelf area and colonized present-day West Sulawesi by marine dispersal; this requires, however, the assumption that the dispersing proto-*Marosatherina* had a tolerance for marine conditions. Such a salt tolerance might indeed be a plesiomorphic character of sailfin silversides.

The occurrence of *Kalyptatherina*, *Neostethus* (Phallostethidae, present with one species on Sulawesi (*Neostethus djajaorum*; Parenti & Louie, 1998), and also some *Pseudomugil* species in brackish habitats, provides support for this assumption. Likewise, *Marosatherina* as well as *Telmatherina bonti* tolerate brackish waters under aquarium conditions (F.H., pers. obs.). However, none of the Sulawesi sailfin silversides has ever been reported from such habitats in nature, and it remains thus unclear, if the assumption that ancestral sailfin silversides were able to cross marine barriers, is in fact realistic. The endemism of all recent species of the family provides a substantial argument against profound abilities for marine dispersal. As an alternative, partly complementary scenario, the ancestral population of *Marosatherina* might have been widespread across both Sunda and Sahul Shelf areas, followed by extinction across large extents during periods of sea-level fluctuations. In that case, *Marosatherina* would represent a relict species as several areas of West Sulawesi remained above sea level or at least were covered only by shallow water during the island's history according to palaeogeographic reconstructions (e.g., Hall, 2009b, 2012b). However, this again assumes that the ancestral population might have been, at least to some degree, saltwater-tolerant.

In line with earlier workers (e.g., Unmack et al., 2013), the present results support the monophyly of the rainbow fishes (Melanotaeniidae). As expected, its genera *Chilatherina* and *Glossolepis* are nested within *Melanotaenia*, and the three major geographic clades recovered correspond to the expected freshwater ecoregions of 'southern', 'northern' and 'western' New Guinea (Abell et al., 2008; Unmack et al., 2013). The island's Central Highlands are the major barrier putatively separating the 'northern' and 'southern' clades, and provide an opportunity to estimate divergence rates also within the Telmatherinidae (see below).

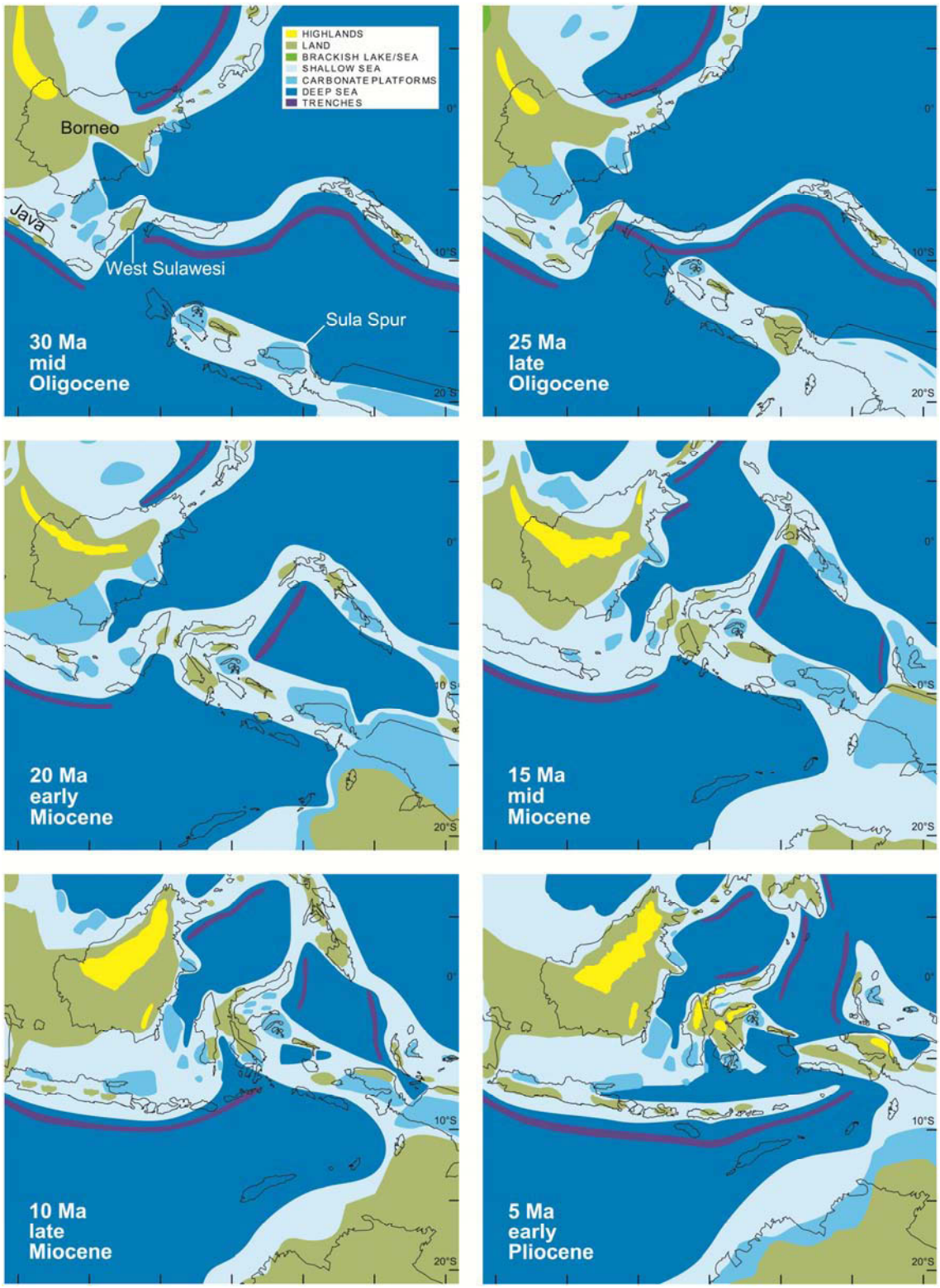


Figure 3.4. Palaeogeographic maps of SE Asia, with particular focus on West Sulawesi and the Sula Spur. Modified from Hall (2012a), with permission (see text for details).

Divergence time estimation: The Malili Lakes radiation

Depending on the method applied, molecular clock approaches estimate the onset of the Malili Lakes radiation to 5.2-17.4 My, but the youngest estimate appears by far the most plausible, given the estimated timeframe for Sulawesi's final amalgamation (see above). This implies ages of about 1.9 My for the initial split of the benthic sharpfins and the predominantly pelagic to benthopelagic roundfins, estimates of c. 1.0 My for divergence within these two lineages inside Lake Matano, and substantially less (0.4 My) for the lineage of haplotypes introgressed by stream populations into Matano's sharpfins.

Roy et al. (2007) provided a first age estimation for the divergence among the three mitochondrial haplotype clades present in Lake Matano's *Telmatherina* radiation. This approach was based on a constant rate of evolution, and applied a genetic distance-age ratio of 1-2%/My. For the sailfin silversides endemic to Lake Matano, the deepest and according to geological data (see Wilson & Moss, 1999 and Robert Hall, pers. comm.) oldest lake of the system, these analyses suggested an age of 0.95-1.9 My separating roundfins ('Clade I' in Roy et al., 2007; see Herder & Schliewen, 2010 for the identity of these clades) and sharpfins ('Clade II'). Haplotypes originating from streams and rivers, present in Lake Matano's sharpfins due to introgressive hybridization (Herder et al., 2006a), diverged from the lacustrine haplotypes (sharpfins + roundfins) in Roy et al.'s (2007) analyses 1.85-3.7 My ago. Their ingroup dataset did however not include the other sailfin silversides species occurring outside of Lake Matano, from the remaining lakes, rivers, and streams of the Malili Lakes; it appears accordingly unclear if this framework is suited for providing reliable estimates for the relevant splits. In the present study, we combine geological and indirect calibration points, as well as recently suggested substitution rates, to estimate and carefully discuss the timing of the most relevant splits within the sailfin silversides radiation in a relaxed molecular clock framework.

Analysis [A] – indirect calibration for the Iso–Melanotaenia *split*

We first used the split between *Iso hawaiiensis* and *Melanotaenia lacustris* as an indirect calibration point, obtained from a study of ricefishes (Adrianichthyidae) by Setiamarga et al. (2009), in analysis [A]. This approach provides an age of 17.4 My for the Malili Lakes radiation (node 15); Lake Matano's roundfins are estimated to an age of 3.7 Mya (node 18) in that analysis (Table 3.2). Under analysis [A], the age of the Malili Lakes radiation (node 15) significantly predates the proposed age for the formation of present-day Sulawesi, namely the final amalgamation of the North and West arms of Sulawesi with the Sula Spur (c. 10-20 Mya; Hall, 2009b; Stelbrink et al., 2012), as well as the geological evidence for the age of the Lake Matano.

Some technical issues might account for the observed inconsistencies between the proposed ages of these nodes, and the geological reconstruction of Sulawesi's history and its

ancient lakes. Dating based on indirect calibration points presupposes that adequate calibration points were used in the source analysis. The 21 fossil and six biogeographic calibration points used by Setiamarga et al. (2009) are remarkable, but all fall outside the atherinomorphs. These priors, plus possible saturation effects, may have led to the bias of an overestimation of splits within the atherinomorphs. A recently published, and more comprehensive dated phylogeny by Near et al. (2013), might justify this inference. There, the atherinomorphs are estimated to have originated c. 80 Mya, while the genera *Iso* and *Melanotaenia* are more recent compared to the split in Setiamarga et al.'s study (Setiamarga et al., 2009: 93-113 My *vs.* Near et al., 2013: c. 50 My); however, some basal nodes remain unresolved in Near et al. (2013), and thus both genera do not form a sister group, as in Setiamarga et al. (2009). Thus, we conclude that the inferred node ages from analysis [A] very likely substantially overestimate the age of the Malili Lakes radiation, as well as that of the intralacustrine speciation and hybridization processes.

Analysis [B] and [C] – indirect calibration for Melanotaenia *clades and rate of 1%/My*

Estimated divergence times for the New Guinean northern and southern *Melanotaenia* clades obtained from Unmack et al.'s study (Unmack et al., 2013) have been used as indirect calibration points in analysis [B]. In addition, we used the substitution rate of 1%/My suggested by Unmack et al. (2013) in their rainbowfish dataset (analysis [C]). Interestingly, analysis [B] shows comparable ages for node 8 and node 9 compared to Unmack et al. (2013), however, the resulting ucld.mean (uncorrelated lognormal relaxed clock mean) is 0.47%/My for this analysis. Accordingly, age estimates inferred from analysis [B] clearly predate the hypothesized starting point of the New Guinean Highlands uplift (node 9: 25.9 My). However, the MRCA of the *Melanotaenia* clade is 39.5 My, and is therefore quite similar to the divergence time estimates of the single *Melanotaenia* lineage in Near et al. (2013).

Likewise, divergence times differ remarkably between the few comparable nodes of analysis [C] and Unmack et al.'s study (Unmack et al., 2013) (see Table 3.2). The considerable mismatch between analysis [C] and the study by Unmack et al. (2013) might be related to the different prior conditions in the respective analyses. With respect to the target question, the taxon sampling available is more complete in the present study. Differences in the resulting estimates may also be due to the genetic markers applied: substitution rates vary across genetic markers and among taxa, and might therefore cause deviating node ages though using the same prior substitution rate.

Analysis [D] – geological calibration using the Central Highlands formation

In a final analysis, we used age estimations for the Central Highlands in New Guinea, and the clades of rainbow fish separated by this major barrier, for estimating the age of the relevant nodes in the sailfin silversides phylogeny (analysis [D]). The Central Highlands in New Guinea

are a massive mountain range, spanning almost the entire island with a maximum altitude of c. 5,000 m above sea level. The continuous formation of the Central Highlands since the late Miocene (c. 10-14 Mya; Hill & Hall, 2003) clearly resulted in the interruption of former watersheds, and led to the present-day regions of freshwater biodiversity. This explains the presence of two well-resolved clades of rainbow fishes north and south of this mountain range (McGuigan, 2000; Unmack et al., 2013), this study); the ancestral populations became most likely continuously separated. However, given the nature of geological events, which should in most cases be considered rather as continuous processes than as distinct, precise events, it appears problematic to reconstruct the exact timeframe in which the proposed ancestral melanotaeniid population in New Guinea was initially separated, giving raise to the two distinct clades observed to date.

Compared to the indirect calibration approach, the prior age of node 9 is substantially younger in analysis [D] (mean age: 10.9 My) than the posterior age inferred from analysis [A] (mean age: 34.5 My). This results in a likewise substantially more recent age of both, the onset of the Malili Lakes radiation, and that of its radiating clades (see Table 3.2 and Fig. 3.3). Accordingly, node ages derived from analysis [D] appear more plausible than those from analyses [A] and [B]. However, the root height representing the age of *Iso* considerably underestimates the divergence time inferred from Setiamarga et al. (2009) by about 70 My, while in contrast, *Iso* is estimated to be c. 15 My younger in analysis [D] compared to the study by Near et al. (2013). This underestimation may not only be due to conflicting topologies, the position of *Iso* within the atherinomorphs (see Discussion above), and the comparatively recent calibration point used (node 9), but might be also correlated with saturation effects in basal nodes by solely using mitochondrial markers.

Telmatherina of Lake Matano's endemic sharpfin radiation carry either mitochondrial haplotypes closely related to those of the lakes' roundfins, or those introgressed by riverine populations (Herder et al., 2006a; Schwarzer et al., 2008; Herder & Schliewen, 2010). The age of the haplotypes introgressed into the Matano flock (mainly node 22) is comparatively young, estimated in analysis [D] to less than 400,000 years. In contrast, age estimates for the 'native' Matano sharpfin haplotypes (node 19: 0.9 My), its sister – the roundfins (node 18: 1.0 My) –, and the clade of *T. celebensis* from the lower lakes of the system (node 26: 0.7 My) are comparatively older, and appear largely congruent to the general age estimates for the Malili Lakes (see above). It appears therefore very likely that the mitochondrial introgression observed has occurred rather recently, in comparison to the age of the lakes' native haplotypes (see Fig. 3.3). Shared haplotypes in highly distinct lake- and stream-dwelling *Telmatherina* are also observed in Lakes Towuti's and Mahalona's *T. celebensis*, and several populations of *T. bonti* (see also Fig. 3.1 and Supplementary Figure 1 in Herder et al., 2006a, incorporating more stream populations). Node 26, comprising these lacustrine and riverine populations, is estimated

to c. 700,000 years ago in analysis [D], a time that coincides with the proposed age for Lake Towuti (Vaillant et al., 2011).

Similarly to the Lake Matano *Telmatherina* radiation, *Paratherina* represents a small, monophyletic radiation, occurring in off- and inshore waters of Lakes Mahalona and Towuti. Analysis [D] suggests that the first diversification event within *Paratherina* (node 27) has occurred c. 1.6 Mya. It appears reasonable to assume that the onset of diversification may have taken place in the older Lake Mahalona, from where the putatively younger Lake Towuti has been colonized. The *Paratherina* populations of both lakes are likely still connected via Tominanga River, as suggested by the recent microsatellite study of Walter et al. (2011). Further support for possible riverine dispersal of *Paratherina* comes from the historical presence of species shared with Lake Towuti in the small hill-lake Lontoa (or Wawontoa; see Kottelat, 1990), connected to the large lake by rivers. This lake has however undergone substantial degradation, and the presence of *Paratherina* could not be confirmed during recent surveys (F.H., pers. obs.).

Conclusions

Divergence times inferred for the Malili Lakes radiation clearly predate both the final formation of Sulawesi and any suggested age estimates for the Malili Lakes for the majority of analyses performed. Hence, node ages derived from analysis [D], based on geological calibration by the New Guinean highland barrier, seem most plausible to us. This means that divergence times obtained from other sources, i.e., Setiamarga et al., (2009); Unmack et al. (2013), possibly overestimate telmatherinid and melanotaeniid clade ages, again highlighting the issues related with molecular clock analyses (see Discussion above).

Some concluding remarks can be made on speciation and hybridization processes in the sailfin silversides radiation, based on the – in our view – most plausible analysis [D]. Based on the present topology, riverine *Telmatherina bonti* populations not only cluster within the lacustrine clades, indicating hybridization events; some of these riverine lineages also appear basal to some lacustrine populations. In line with similar results from an earlier phylogenetic study (Herder et al., 2006a), this clearly indicates that the Malili Lakes were colonized by riverine populations, which appears highly plausible, and meets patterns observed also in invertebrate radiations, like e.g., the pachychilid snails (von Rintelen et al., 2004).

The present study indicates that the Sulawesi telmatherinids might have originated c. 3-5 Mya, a period when present-day Sulawesi was being formed through a series of tectonic events such as e.g., mountain uplifts in West Sulawesi and the Matano fault formation (Wilson & Moss, 1999; Hall, 2012b). The Matano fault gave rise to the rift lake Matano, and probably also initiated the formation of the remaining lakes of the Malili Lakes system. Preliminary

geological reconstructions suggest that the Malili Lakes are generally 1-2 My old, while preliminary seismic data support with 600,000-700,000 years a younger age for Lake Towuti. This geological and seismic evidence is in line with the present phylogenetic reconstruction, which shows comparatively recent diversification and hybridization events within *Telmatherina celebensis* and their riverine relatives, compared to diversification and introgression in Lake Matano (node 26; see Fig. 3.1 and 3.3). Age estimates suggest diversification along a benthic-pelagic axis, into sharpfins and roundfins, c. 1.9 Mya after Lake Matano was colonized by stream populations, followed by a rapid radiation in both of these clades in the last 1 My. Secondary hybridization did probably not affect initial divergence within Lake Matano's sharpfin radiation, as the age of the introgressed haplotypes clearly postdates the initial diversification by about 600,000 years.

Chapter 4

A snail perspective on the biogeography of Sulawesi, Indonesia: origin and intra-island dispersal of the viviparous freshwater gastropod *Tylomelania*

Introduction

The Indonesian island Sulawesi lies at the heart of the Indo-Australian Archipelago (IAA), one of the world's most species-rich areas and arguably the one with the most complex geography and geology (Lohman et al., 2011). The largely insular nature of the region sets it apart from the other major tropical rainforest areas (Morley, 2000). While the entire IAA is essentially the product of continental collisions since the Paleozoic (Metcalfe, 2011b), the eastern part of the archipelago has its origin in the ongoing collision between Asia and Australia since the early Miocene (Hall, 2011). An ancient continental core in the West that has been emergent since the Mesozoic – comprising Indochina and Sundaland – is thus sharply set apart from an area of predominantly young (Miocene to Pleistocene) oceanic islands including Sulawesi in the East bordered by the Australian continental margin (Hall, 2012b). The changes in the palaeogeography of the IAA have left a strong signature in the distribution of the region's fauna and flora. Several faunal boundaries, including the famous 'Wallace Line' (Wallace, 1860), were proposed to account for the observation that the western or eastern distribution limits of many animal groups are located within the oceanic island region of the IAA (see e.g., Simpson, 1977). The transitional region Wallacea, which comprises all islands between the Sunda and Australian shelves, was proposed as one attempt to avoid the designation of a single faunal boundary (Dickerson, 1928). While the concept of a transitional zone was quickly shown to be inadequate in itself (Mayr, 1944; Simpson, 1977), the name Wallacea has stuck for the Indonesian part of this oceanic island assemblage (Cox, 2001; Lohman et al., 2011) and was also used to circumscribe one of SE Asia's four biodiversity hotspots (Myers et al., 2000), characterized by its unique, yet highly threatened biotic diversity. The fauna of the Wallacean islands is generally characteristic of oceanic islands (Whittaker & Fernández-Palacios, 2007), e.g., being 'disharmonic", i.e., depauperate at a higher taxonomic level, and having a high proportion of endemic species (Mayr, 1944; Keast, 2001).

Sulawesi stands out among the islands of Wallacea due to its size, different geological history and higher faunal diversity (Whitten et al., 2002). The geographic origin of Sulawesi's fauna and the mode of colonization of the island is still disputed to some degree even 150 years after the first proposal of the Wallace Line (Michaux, 2010; Stelbrink et al., 2012). One reason for the controversy lies in the fact that Sulawesi, in contrast to most of the other islands in Wallacea, is not truly oceanic, but a composite island at the centre of the Asia-Australia collision zone. Parts of the island were formerly attached to either the Asian or Australian continental margin and separated from these areas by vicariant processes (for details see Stelbrink et al., 2012). In the West, the opening of the Makassar Strait separated West Sulawesi from Sundaland in the Eocene c. 45 Mya (Hall, 2011). In the East, the traditional view of collisions of multiple micro-continental fragments sliced from New Guinea with an active

volcanic margin in West Sulawesi at different times since the Early Miocene c. 20 Mya (Burrett et al., 1991; Wilson & Moss, 1999) has recently been replaced with assuming extensional fragmentation following a single Miocene collision of the Sula Spur with West Sulawesi (Spakman & Hall, 2010; Hall, 2011; Stelbrink et al., 2012). However, the revised geological hypothesis does not alter the theoretical possibility to invoke vicariance-based hypotheses for an origin of taxa on Sulawesi from either Asia or Australia (Stelbrink et al., 2012).

The predominantly Asian origin of the island's biota has been recognized for a long time and has during the last two decades gained support from molecular phylogenetic studies involving taxa from across the Wallace line (see reviews by Lohman et al., 2011 and de Bruyn et al., 2012). The recent dating of the colonization of Sulawesi in 20 non-marine animal groups by a molecular clock approach has shown that Miocene to Pleistocene dispersal to the island from Asia (Sundaland, Philippines) is the most likely mechanism for the origin of the vast majority of Sulawesi taxa (Stelbrink et al., 2012). However, an origin of Sulawesi taxa from both Asia and Australia (including New Guinea) through vicariant processes could not be ruled out in some instances, e.g., from Asia for mite harvestmen (Clouse & Giribet, 2010) or from Australia for pachychilid freshwater snails (Glaubrecht & von Rintelen, 2003), atherinimorph fishes (Sparks & Smith, 2004), and phalangerids (Ruedas & Morales, 2005).

Among the three taxa with an Australian origin, the pachychilid gastropods likely represent the strongest case for a vicariance scenario as the likelihood for dispersal in these strictly freshwater-dwelling animals seems extremely low. On Sulawesi, the group is solely represented by the endemic genus *Tylomelania* Sarasin & Sarasin, 1897 (von Rintelen & Glaubrecht, 2005). *Pseudopotamis* Martens, 1894, restricted to two of the North Australian Torres Strait islands, has been consistently identified as the sister group of *Tylomelania* (Glaubrecht & von Rintelen, 2003; Köhler et al., 2004; Köhler & Dames, 2009; Köhler & Glaubrecht, 2010). Both taxa are ovoviviparous and share the synapomorphy of a pallial oviduct modified into an uterine brood pouch releasing comparatively large and shelled juveniles (von Rintelen & Glaubrecht, 2005). Brooding in freshwater snails is generally regarded as being associated with a low dispersal potential (Cohen & Johnston, 1987; Glaubrecht, 1996) which renders transoceanic dispersal across a distance of c. 2,000 km between Sulawesi and the Torres Strait Islands unlikely. This has prompted a hypothesis of vicariance through 'terrane rafting' from the north Australian margin (Glaubrecht & von Rintelen, 2003; von Rintelen & Glaubrecht, 2005). Again, the recent replacement of the concept of terrane rafting by the assumption of extensional fragmentation has not altered the basic premise of this hypothesis (see Stelbrink et al., 2012). While the study of Stelbrink et al. (2012) has demonstrated that the estimated timing of the split between *Tylomelania* and *Pseudopotamis* is consistent with the geological data under the 'out-of-Australia' vicariance hypothesis, only a single sequence each from two species of *Tylomelania* was included.

As a result of its complex geological history (compare above), Sulawesi is geographically highly structured, primarily through its subdivision into the four arms constituting its k-shape. In addition, various mountain ranges, mostly resulting from the amalgamation of its constituent fragments in the Miocene (Hall, 2012b), and some low-lying areas (e.g., Tempe and Gorontalo depressions) inundated during sea-level highstands act as potential barriers to dispersal across the island. This is reflected in the distribution patterns of species on Sulawesi, which exhibit strong geographic structuring in all taxa examined where sampling coverage across the island is sufficiently dense (Butlin et al., 1998; Evans et al., 1999, 2003c, 2008; Bridle et al., 2001; Setiadi et al., 2011; Linkem et al., 2013). For some taxa, the phylogeographic breaks or contact zones, respectively, were found to be largely congruent, suggesting a strong role for habitat fragmentation in the diversification of Sulawesi taxa (Evans et al., 2003c, 2008). For freshwater organisms, respective comparative data are still lacking, as most research has focused on the endemic aquatic radiations in the ancient lakes of Sulawesi (see review in von Rintelen et al., 2012). *Tylomelania* should serve as a good model to trace the effect of historic vicariant barriers as the animals are not only restricted to freshwater, but their present-day distribution suggests that they are not capable of occurring at an altitude of more than about 700 m and they are only found near the coast where rivers or streams are directly descending from nearby mountains, or in karstic outcrops (compare Fig. 4.1 and 4.3, pers. observation 1999-2011).

Consequently, we here study the phylogeography of *Tylomelania* based on samples from across its entire distribution range on Sulawesi to address questions pertaining to the origin of the taxon on Sulawesi and its subsequent diversification on the island. Specifically, we attempt to identify an area of origin on Sulawesi from the sequence of splits within *Tylomelania* and to link patterns of intra-island differentiation to Sulawesi's geology and topography. In addition, we discuss our data in comparison with those derived from other taxa that are widely distributed on Sulawesi, such as toads (Evans et al., 2003c, 2008), macaques (Evans et al., 2003b, 2003c), fanged frogs (Setiadi et al., 2011), and tarsiers (Merker et al., 2009).

Materials and Methods

Material

Samples from 191 sites comprising 1,170 individuals of 62 species including 26 undescribed morphospecies (species delimitations are based on shell and radula characters that have been shown to be effective in distinguishing sympatric species in *Tylomelania* from the lakes of Sulawesi; von Rintelen et al., 2007) were collected across the entire distribution range of *Tylomelania* on Sulawesi (Fig. 4.1, Table S1). All material has been preserved in 70-95% ethanol. Voucher specimens employed in this study are deposited in the Malacological

Department and DNA samples are stored in the central DNA collection of the Museum of Natural History, Berlin (ZMB). Locality details and both museum and EMBL accession numbers for all sequenced animals are provided in Table S1 and S2.

Molecular methods

For the phylogeny, two mitochondrial gene fragments, a ~890 bp region of the 16S ribosomal RNA gene (16S) and a 710 bp (the so-called DNA barcoding fragment) fragment of the Cytochrome Oxidase subunit I gene (COI) were sequenced in 153 specimens of *Tylomelania* using methods and primers described previously (von Rintelen et al., 2004; von Rintelen & Glaubrecht, 2008). The dataset was complemented by published sequences of both genes from 42 individuals (von Rintelen et al., 2004). Sequences of the two species of *Pseudopotamis* from the North Australian Torres Strait Islands were used as outgroup as suggested by previous studies (Glaubrecht & von Rintelen, 2003; Köhler et al., 2004; Köhler & Glaubrecht, 2010). COI was sequenced from an additional 968 specimens for the haplotype networks.

Phylogenetic analyses

Orthologous DNA sequences were aligned by eye (COI) and with MUSCLE (16S) (Edgar, 2004). The MUSCLE alignment was corrected manually for unambiguous algorithm-specific errors. Following alignment, the 202 sequences obtained for each gene were reduced to unique haplotypes (Table S1) using DAMBE v. 5.066 (Xia & Xie, 2001). The aligned sequence sets of COI (660 bp, 127 sequences; see Fig. S1) and 16S (868 bp, 136 sequences; see Fig. S2) were analyzed separately and combined into a single concatenated alignment of 157 sequences (1,528 bp; see Fig. 4.1).

Maximum Parsimony (MP) analyses were performed with PAUP* v. 4.0b010 (Swofford, 2002); bootstrap replicates = 10,000; gaps were treated as fifth state). For Bayesian Inference (BI) analyses, the appropriate models of sequence evolution (HKY+G for COI and GTR+I+G for 16S) was determined using jModelTest v. 0.1.1 (Posada, 2008) both based on the Akaike Information Criterion and the Bayesian Information Criterion. The two genes were set as partitions in the concatenated dataset and analyses run with the model specified for each partition separately: ML with RAxML BlackBox (Stamatakis et al., 2008) using GTR+G and 100 bootstrap replicates, and BI with MrBayes v. 3.1.2 (Ronquist & Huelsenbeck, 2003); ngen = 5,000,000 for single-gene datasets, and 10,000,000 for the combined dataset; samplefreq = 100 / 200; burnin = 35,001). For comparison, trees based on analysis of the separate datasets are shown in the supplementary information (Fig. S1 and S2).

Molecular clock analyses

A reduced COI dataset (*n* = 26) was tested for nucleotide substitution saturation using the test by Xia et al. (2003) in DAMBE showing no significant saturation for COI. Strict and uncorrelated lognormal relaxed clock analyses were performed in BEAST v. 1.7.3 (Drummond & Rambaut, 2007); GTR+I+G; Yule process; ngen = 50,000,000; log = 1,000; burnin = 35,001) using a substitution rate of 1.76%/My as suggested by Wilke et al. (2009) for several freshwater taxa including gastropods. Fossils were not available for calibration. Log files of both runs were subjected to a Bayes factor analysis as implemented in Tracer v. 1.5 (Rambaut & Drummond, 2007) resulting in a small value of 0.212 slightly supporting the strict clock analysis.

Haplotype networks

The complete COI dataset (1,170 sequences) was reduced to 233 unique haplotypes (Fig. S3 and S4, Table S2) using DnaSP v. 5.10.01 (Librado & Rozas, 2009). Haplotype networks were calculated from these haplotypes using TCS v. 1.2.1 (Clement et al., 2000) using the 95% parsimony criterion.

Results

Phylogeny of Tylomelania

The molecular phylogeny based on MP, ML and BI analyses of the concatenated mtDNA dataset (COI and 16S) reveals eight major lineages (six clades and two distinct haplotypes) of *Tylomelania* on Sulawesi (Fig. 4.1-3; Table 4.1). The monophyly of the six clades is supported by Bayesian Posterior Probabilities (BPP) >0.9, while only four have a MP and ML bootstrap support >80%. The splitting sequence of the eight lineages is not supported at all in the MrBayes topology (BPP < 0.9; Fig. 4.1 and 4.3) and only partially in the BEAST topology (Fig. 4.2). Separate analyses for COI and 16S yielded essentially the same topology (Fig. S1 and S2), differing only in the level of support for the major lineages. The eight major lineages are quite distinct, with an inter-lineage genetic distance range of 5.2-12.1% (COI) and 2.6-8.4% (16S), respectively (Table 4.2). A strong geographic pattern is evident in the distribution of the lineages of *Tylomelania*, with no to limited overlap in their distribution ranges (Fig. 4.2 and 4.3). Three clades (1,3,4 – 'red', 'blue', and 'green' in Fig. 4.2 and 4.3) are more widely distributed in East Central Sulawesi (clade 1), South-South West, South Central, and Southeast Sulawesi (clade 3), and Southwest Sulawesi (clade 4). All other lineages are confined to comparatively small regions. The two basal lineages of *Tylomelania* (forming sister groups in the BEAST analysis, Fig. 4.2) both occur on the easternmost Luwuk peninsula of the eastern arm of Sulawesi. The relationship of these two lineages to each other is not supported (BPP < 0.7) in either BI analysis (Fig. 4.1 and 4.2), though.

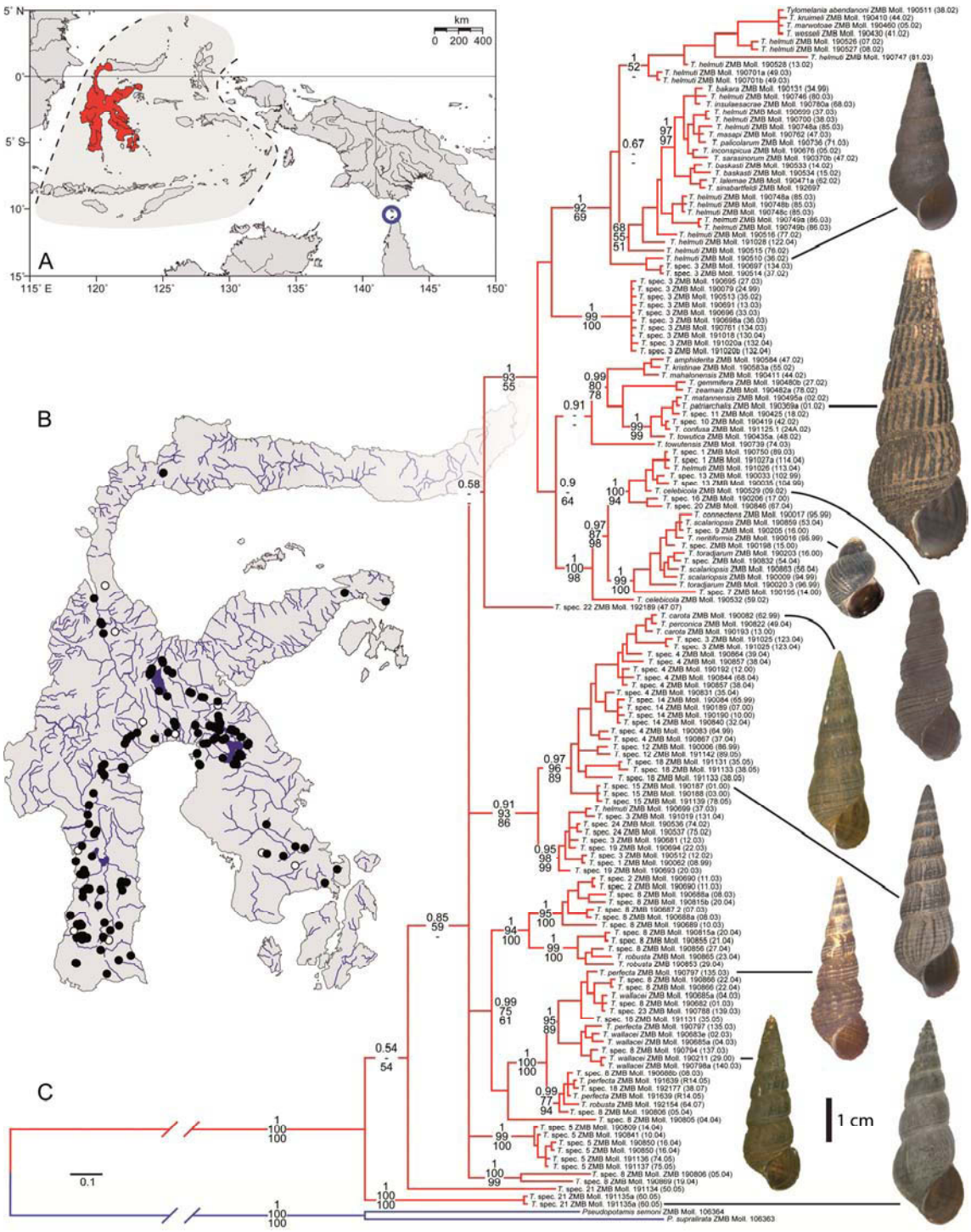


Figure 4.1. Distribution and molecular phylogeny of *Tylomelania*. A. Distribution area of *Tylomelania* (red) and its sister taxon *Pseudopotamis* (blue circle) on Sulawesi and the Torres Strait Islands, respectively. B. Map of Sulawesi with sample sites of *Tylomelania*. Black dot – sample sequenced for this study; white dot – museum sample. C. BI phylogram and shells of selected species of *Tylomelania*. Phylogeny based on 1,528 bp of mtDNA (COI and 16S). Numbers on branches show node support; BI posterior probability (top), ML (centre), and MP (bottom) bootstrap values. Colour code as in panel A.

Table 4.1. Intra- and inter-lineage genetic distances (p-distance) in *Tylomelania*. Top-right cells: COI; bottom-left cells: 16S. Intra-clade genetic distances are shown in brackets.

	Clade 1 (0.0-9.1)	**Clade 2 (-)**	**Clade 3 (0.2-4.4)**	**Clade 4 (0.2-7.6)**	**Clade 5 (0.2-0.8)**	**Clade 6 (3.6)**	**Clade 7 (-)**	**Clade 8 (-)**
Clade 1 (0.0-5.7)	-	6.3-8.3	5.6-10.5	6.0-11.1	5.8-9.4	7.1-10.5	7.9-10.6	9.8-12.1
Clade 2 (-)	3.2-5.7	-	6.0-7.7	5.2-6.9	5.6-6.0	6.6	8.5	8.7
Clade 3 (0.0-4.3)	3.4-7.4	3.3-4.9	-	5.5-9.2	5.6-7.6	6.1-7.7	8.7-10.3	9.0-10.8
Clade 4 (0.0-4.1)	2.7-6.4	3.1-4.2	3.2-5.7	-	5.2-7.4	6.0-8.4	8.1-9.7	9.0-10.6
Clade 5 (0.1-0.2)	3.3-6.0	2.6-2.7	3.1-4.8	2.7-3.9	-	6.5-7.3	8.1-8.7	9.5-10.0
Clade 6 (3.0)	4.2-7.0	3.7-4.2	3.6-5.4	3.7-4.9	3.3-3.9	-	9.8-10.0	10.5-11.3
Clade 7 (-)	6.0-8.4	5.9	5.4-7.1	5.4-6.6	5.8-5.9	5.9-6.2	-	9.2
Clade 8 (0.1)	5.9-8.2	5.8-5.9	5.4-7.0	5.3-6.6	5.0-5.3	6.0-6.4	6.3-6.4	-

Divergence time estimates

The results of the BEAST strict molecular clock analysis (Fig. 4.2A, Table 4.2) suggest that the split between *Tylomelania* from Sulawesi and *Pseudopotamis* from the North Australian Torres Strait Islands occurred in the mid Miocene c. 19.5 Mya (Fig. 4.2, Table 4.2, node 1) and the first diversification event on Sulawesi at the Miocene-Pliocene transition c. 5.4 Mya (Fig. 4.2, Table 4.2, node 2). This basal split between the Luwuk peninsula lineage(s) and the rest of *Tylomelania* is followed by the subsequent diversification of the latter lineage into the major clades of *Tylomelania* in the mid to late Pliocene (c. 3-4 Mya).

Table 4.2. Estimated maximum mean node ages from the strict clock analysis for COI (see Fig. 4.2). HPD – high posterior density.

Node no.	**Mean (My)**	**95% HPD min.**	**95% HPD max.**
1	19.51	12.48	27.37
2	5.37	4.16	6.59
3	4.48	3.18	5.78
4	4.07	3.32	4.86
5	3.57	2.87	4.30
6	3.10	2.45	3.83
7	2.95	N.A.	N.A.
8	2.54	1.87	3.22
9	2.51	1.88	3.15

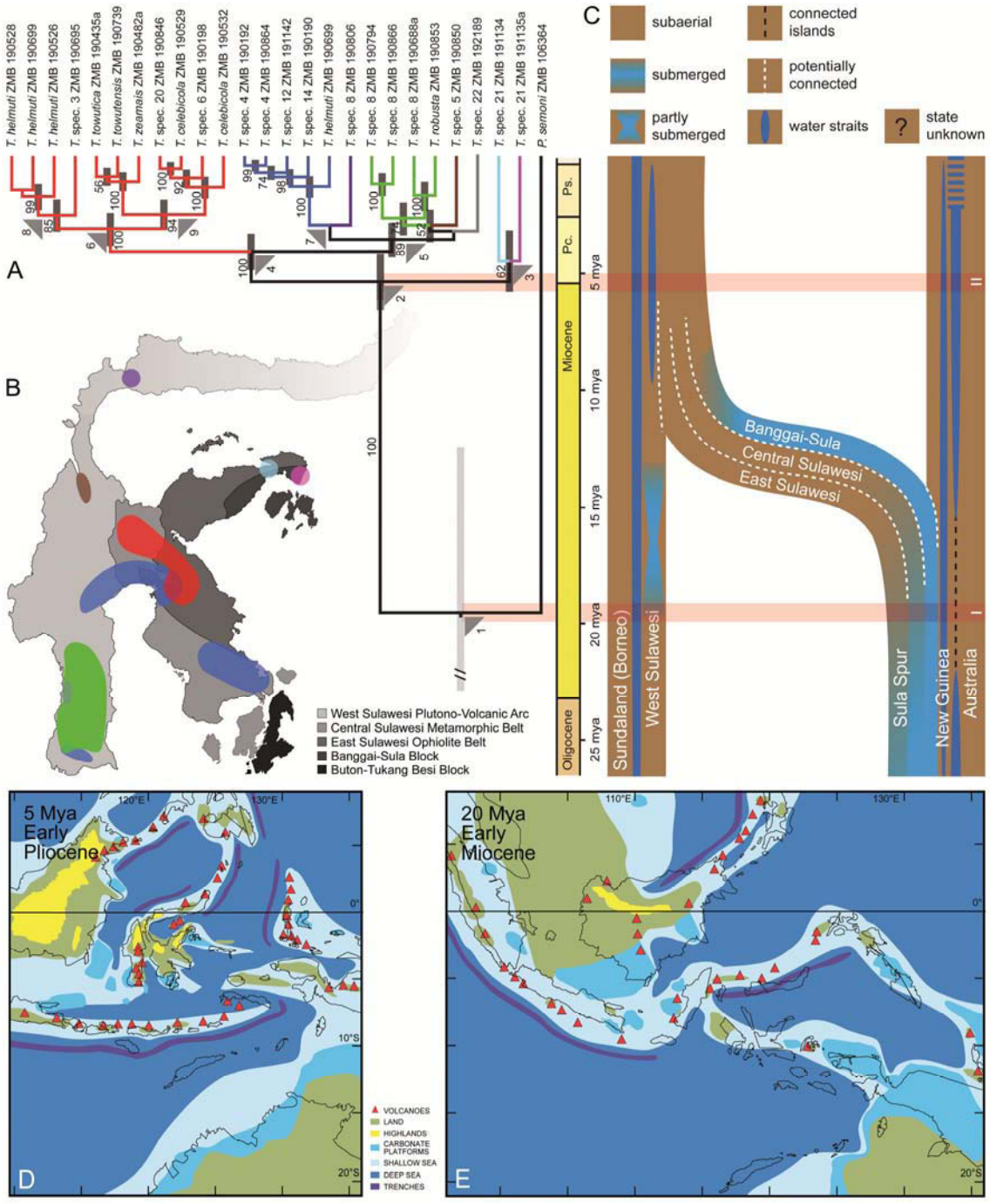


Figure 4.2. Calibrated phylogeny of *Tylomelania* and the tectonic history of Sulawesi. A. Bayesian (BEAST) chronogram of *Tylomelania* based on 660 bp of mtDNA (COI) using a substitution rate of 1.76%/My (see Materials and methods). The major lineages of *Tylomelania* are colour-coded. The pink horizontal bars link the events associated with nodes 1 (TMRCA of *Tylomelania* and *Pseudopotamis*) and 2 (first speciation event within *Tylomelania*) with the geological timescale and palaeogeography of Sulawesi (panel C). B. Map of Sulawesi with major tectonic subdivisions (compare panel C) and the distribution of the major lineages of *Tylomelania* (colour-coding of areas corresponds to major lineages in panel A). C. Schematic summary of the geographic connections and the timing of separation or collision of the different parts of Sulawesi (see map in panel B). Blue tinting indicates presumably submerged areas, see legend for details. Modified from Stelbrink et al. (2012). D,E. Palaeogeographic reconstructions of the Asia-Australia collision zone 5 Mya and 20 Mya. Modified from Hall (2009b).

Spatial genetic diversity within Tylomelania

The TCS analyses yielded 26 distinct parsimony haplotype networks under a 95% confidence interval cut-off threshold (Table 4.3 and Table S1). Matching these haplotype networks to the major clades or lineages, respectively, obtained in the analysis of the concatenated dataset reveals that a strong geographic structure in the genetic diversity of *Tylomelania* below the level of the major clades as well (Fig. 4.3). In two areas in Southwest Sulawesi and in Central Sulawesi, two each of the three largest clades in terms of known distribution area (clade 1,3,4; red, blue and green in Fig. 4.2 and 4.3) of *Tylomelania* overlap and haplotypes from both respective clades are found within a single population.

Discussion

The origin of Tylomelania *on Sulawesi*

The results of this study are compatible with the 'out-of-Australia' vicariance hypothesis for the origin of *Tylomelania* on Sulawesi (Glaubrecht & von Rintelen, 2003; von Rintelen & Glaubrecht, 2005; Stelbrink et al., 2012). The split of *Tylomelania* and *Pseudopotamis* (c. 19.5 Mya) matches or predates the separation of the Sula Spur from mainland New Guinea at 18-13 Mya (Fig. 4.2, horizontal red bar marked 'I') and the first split within *Tylomelania* (c. 5.4 Mya) on Sulawesi matches or postdates the fusion of the different parts of Sulawesi at 10-6 Mya (Fig. 4.2, horizontal red bar marked 'II'). Our results are largely in line with those from (Stelbrink et al., 2012), who suggested an older age of 30.5 Mya (95% HPD: 42-20 Mya) for the split of *Tylomelania* and *Pseudopotamis* and a slightly younger age of c. 4.7 Mya (95% HPD: 1.9-7.8 Mya) for the first diversification event within *Tylomelania*. The congruence (considering the wide overlap in 95% HPD) between the results of both studies seems remarkable given that the dates from (Stelbrink et al., 2012) were based on a much smaller sequence dataset for *Tylomelania* (from Köhler & Glaubrecht, 2010) and a different marker (16S).

As pointed out by Stelbrink et al. (2012), there is a major conflict between the biological evidence as interpreted under the vicariance scenario and the distribution of land according to the geology-based palaeogeographic reconstructions of the Sula Spur area during the Miocene (see e.g., Fig. 4.2 and Hall, 2012b). However, given the difficulties of recovering evidence for land in a region of small islands (Hall, 2009b), we would follow Stelbrink et al. (2012) in arguing that the biological data might actually constitute harder evidence in this case. Under the vicariance scenario, the c. 15 My gap between the split of *Tylomelania* from *Pseudopotamis* and the first diversification event within *Tylomelania* could be explained by the restriction of the ancestral lineage of *Tylomelania* to a small and fluctuating distribution area on the Sula spur prior to the expansion of its distribution range on Sulawesi. These Sula Spur populations have most likely become extinct, as extensive searches on the Banggai (Banggai,

Peleng; 2005, 2008) and Sula islands (Taliabu, Mangole, Sanana; 2008, 2013) were not successful. One aim of this study has been to attempt to identify the region of Sulawesi that was initially colonized by the ancestors of *Tylomelania* by looking at the island-wide phylogeography of the taxon. The topology of the molecular phylogeny (Fig. 4.1-3) suggests that the basal split within *Tylomelania* occurred between populations on the easternmost Luwuk peninsula and the rest of island. The Luwuk peninsula is a fusion zone between East Sulawesi and the westernmost part of the Banggai-Sula fragment, both of which formed part of the Sula Spur (see Fig. 4.2D,E). The amalgamation process of Sulawesi in the wake of the Sula Spur collision with West Sulawesi is poorly understood (Hall, 2012b), but the Banggai-Sula fragment apparently became connected rather late in the late Miocene/early Pliocene to the core of present day Sulawesi, in contrast to parts of East Central and Southeast Sulawesi as borne out by evidence for land there from the Miocene (Hall, 2012b). Against this background, it is tempting to speculate that Sulawesi was colonized by the ancestor of *Tylomelania* from the Banggai-Sula area, particularly given the timing of the split between the lineages of *Tylomelania* from the Luwuk peninsula lineages and the remainder of the island, which at c. 5.4 Mya nicely matches the fusion of that part of the island to the rest of Sulawesi in the late Miocene-early Pliocene (Garrard et al., 1988). The topology of the BI tree as reconstructed with MrBayes (Fig. 4.1) with a paraphyletic position of Luwuk peninsula *Tylomelania* relative to all other species would support this notion. However, the relationship of the Luwuk lineages to each other is not resolved and if the sister group relationship suggested in the BEAST analysis (Fig. 4.2) should be correct rather than the topology derived from the MrBayes analyses (Fig. 4.1), the assumption of an early isolation of a population on the eastern arm of Sulawesi would be an equally parsimonious alternative interpretation of that topology. Further progress on this issue will likely depend on filling the sampling gap between Central Sulawesi and the Luwuk peninsula on the Eastern peninsula.

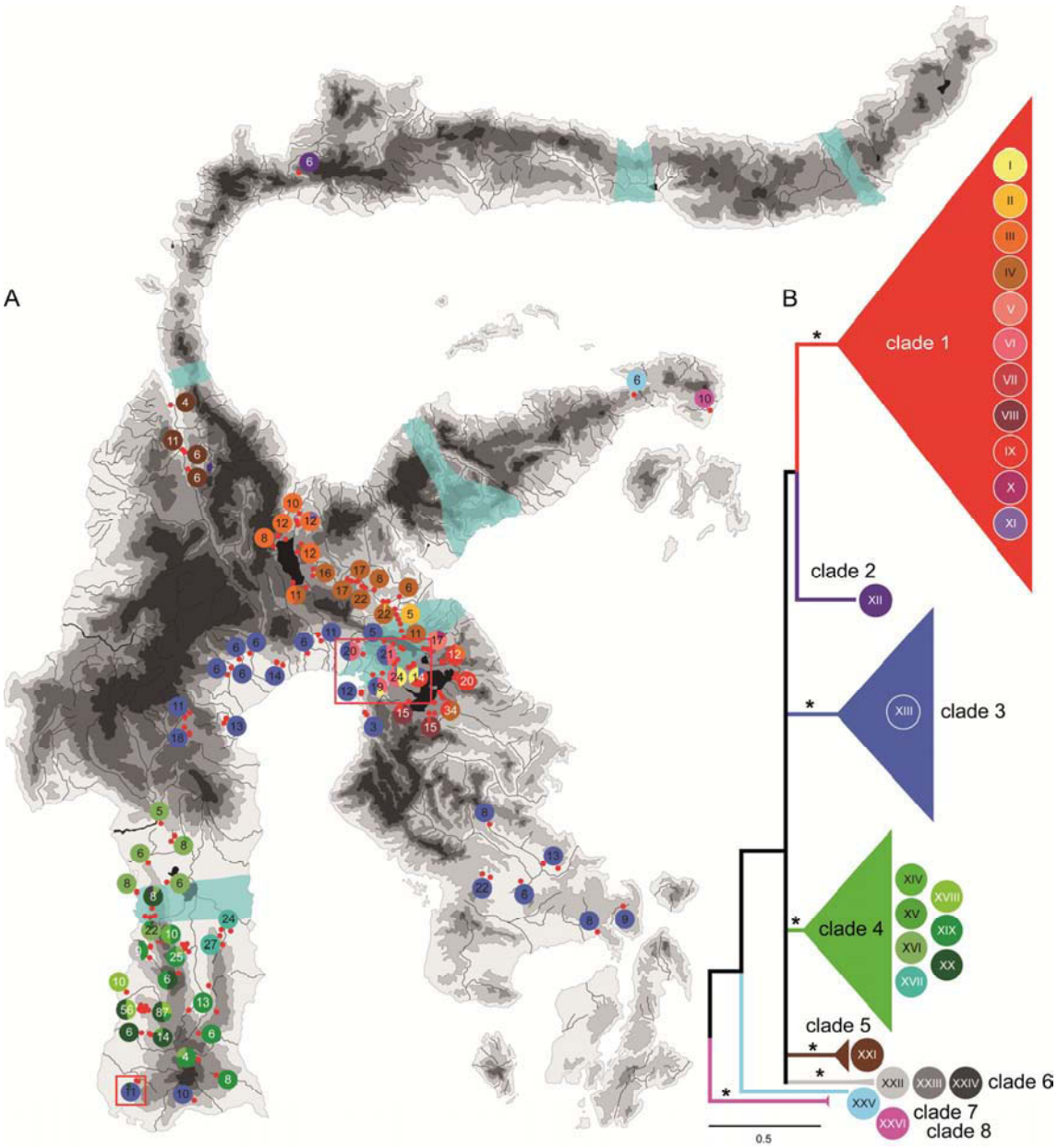


Figure 4.3. Distribution of the major lineages and haplotype groups of *Tylomelania* on Sulawesi. A. Sulawesi map with assignment of sampling sites (small red dots) to haplotype groups (pie charts, haplotype group specific colour code as in panel B). Haplotype group diversity for sites in close geographic vicinity has been subsumed within a single pie chart. Numbers within pie charts are sequenced individuals. Bluish areas indicate the position of contact zones for macaque species on Sulawesi (from Evans et al., 2003c). Red frames show the position of the two contact zones of *Tylomelania* clades. B. BI phylogram based on tree shown in Fig. 4.1, all clades with a BPP of > 0.90 are shown as cartoons. Colours correspond to those used in Fig. 4.2 and the circles with roman numbers indicate haplotype groups within each lineage (see Fig. S1).

Table 4.3. Haplotype networks (groups) of *Tylomelania* with lineage assignment and intra-network genetic distance range (p-distance) for COI.

Lineage/ clade	Haplotype network	N individuals	N haplotypes	COI
1				
	I	30	6	0.2-2.0
	II	7	3	0.6-0.8
	III	74	22	0.2-1.4
	IV	203	40	0.2-2.6
	V	22	4	0.2-0.6
	VI	50	9	0.2-0.6
	VII	12	1	-
	VIII	14	1	-
	IX	46	14	0.2-3.0
	X	3	1	-
	XI	1	1	-
2				
	XII	6	2	0.2
3				
	XIII	282	63	0.2-4.4
4				
	XIV	5	1	-
	XV	2	2	0.5
	XVI	56	7	0.2-1.7
	XVII	53	10	0.2-3.0
	XVIII	61	14	0.0-4.0
	XIX	82	5	0.5-1.1
	XX	111	18	0.0-3.7
5				
	XXI	27	4	0.2-0.8
6				
	XXII	5	1	-
	XXIII	2	1	-
	XXIV	1	1	-
7				
	XXV	6	2	0.3
8				
	XXVI	10	1	-

The geography of diversification on Sulawesi

The major lineages of *Tylomelania* on Sulawesi show a largely allo- or parapatric distribution (Fig. 4.2 and 4.3), while with one exception each lineage has a continuous distribution area. The most widely distributed clade 3 (Fig. 4.2 and 4.3, blue) occurs in three apparently disjunct areas in Southwest, Central and Southeast Sulawesi. The distribution of this clade overlaps with that of the other two widespread clades 1 and 4 (Fig. 4.2 and 4.3, red and green) in South-Southwest Sulawesi and Central Sulawesi, respectively (Fig. 4.3, red frames). The fact that haplotypes assigned to different clades are found among individuals within the same population (same sampling site within a 10-100 m stretch of stream or river, same morphospecies) suggests mitochondrial introgression in these contact zones. Given the distinctness of the mtDNA lineages involved (COI p-distance range between each of the three lineages: 5.6-11.1%), this is somewhat surprising. The lack of data from nuclear markers prevents an in-depth discussion at present, but the apparent absence of morphological hybrids might indicate that the introgression between the lineages is not a recent phenomenon. Different haplotypes are also found among

individuals of a single and morphologically uniform population at the level of haplotype groups (separate networks under the 95% cut-off threshold) within the same clade (Fig. 4.3). At the other extreme of the spectrum, the same haplotype can be shared among morphologically different and sympatrically occurring species in several localities (see e.g., haplotypes 1 or 2 from Southwest Sulawesi in Table S2). Without nuclear data, the cause for these phenomena – e.g., hybridization *vs.* incomplete lineage sorting (see e.g., Funk & Omland, 2003) for possible causes of genetic discordance at the species level) – must remain speculative.

The strong geographic structure of genetic diversity in *Tylomelania* might be explained by various factors such as habitat fragmentation in the widest sense (e.g., vicariance through the rise of mountains or the formation of sea barriers), the development of ecological barriers, isolation by distance, or most likely a combination of these. If habitat fragmentation would be the predominant factor in shaping the distribution of the major lineages in *Tylomelania*, a strong correlation of lineage-specific distribution boundaries with geographic barriers irrespective of distance would be expected. While seemingly trivial, this concept can be hard to test, as today's barriers to dispersal may be very different from those in the past. This is particularly true for *Tylomelania* on Sulawesi, as the diversification into major lineages took place in the Pliocene and coincides with the onset of pronounced orogeny in many parts of Sulawesi (Hall, 2012b). The same is essentially true for ecological factors, such as climate, vegetation etc. However, *Tylomelania* has rather uniform ecological preferences with the notable exception of the species flocks in the ancient lakes of Sulawesi (Glaubrecht & von Rintelen, 2008; von Rintelen et al., 2010, 2012). The homogenous morphology of the rasping tongue (radula), which is indicative of substrate in *Tylomelania*, reflects this (Glaubrecht & von Rintelen, 2008; von Rintelen et al., 2010). Thus, it does not seem very likely that ecological fragmentation has been a major factor in shaping distribution patterns. As geological vicariance will not just affect single taxa but a wide range of organisms, the search for congruent distribution patterns among unrelated taxa is a promising approach to identify such common barriers and a pivotal principle of comparative phylogeography and biogeography (see e.g., Arbogast & Kenagy, 2001; Avise, 2004). On Sulawesi, the species or subspecies of many non-flying terrestrial and limnic taxa that were sampled across the island or larger parts are distributed parapatrically or allopatrically (see Evans et al., 2003c for an overview). Molecular data are only available for a few widely sampled taxa from Sulawesi to date, though. Based on the spatial structure of genetic diversity in Celebes toads (*Bufo celebensis*) and macaques (*Macaca* spp.), Evans et al. (2003c) suggested that their contact zones were congruent and likely reflect habitat fragmentation by physical barriers.

While this interpretation of their data was questioned and isolation by distance proposed as an alternative hypothesis (Bridle et al., 2004), coalescent analyses support the original habitat fragmentation scenario (Evans et al., 2008). Similar patterns were observed in

the fanged frog *Limnonectes* spp. (Setiadi et al., 2011), while *Chitaura* grasshoppers show at least partly deviant patterns (Bridle et al., 2001) and *Lamprolepis* skinks show a very different North-South pattern (Linkem et al., 2013). The primary diversification events of these taxa took place from the earliest Pliocene (fanged frogs) through the middle Pliocene (macaques, grasshoppers) to the late Pliocene/Pleistocene (Celebes toads) (Stelbrink et al., 2012). Assuming that these estimates are correct (see Stelbrink et al., 2012 for some dating issues), the time of the primary speciation events on Sulawesi does not seem to play a major role in determining the congruence of contact zones or areas of endemism. The split into the major lineages of *Tylomelania* is hypothesized to have occurred in the mid Pliocene (Fig. 4.2A) and does match the general timeframe of diversification of the taxa investigated so far. However, neither the distribution of the major lineages of *Tylomelania* nor of the constituent haplotype groups seems to be universally constrained by the barriers associated with the six macaque contact zones (Evans et al., 2003c; Fig. 4.3). While the distribution boundaries of *Tylomelania* clades apparently match three of these, two other contact zones, across the north-central part of the Southwest peninsula and the northern part of the Southeast peninsula of the island, do not form a barrier for the three clades of *Tylomelania* (1,3,4) whose distribution extends across them. This suggests a partial fit of areas of endemism (at different taxonomic levels, though: species groups (lineages) in *Tylomelania* and species or subspecies in macaques or toads) of *Tylomelania* and the other taxa, for which an explanation might be sought through looking at potential differences in the nature (e.g., age) of the respective barriers. However, the sampling of *Tylomelania* is only sufficiently dense, i.e. from the contact zone and both adjoining areas, to make a firm statement for the two contact zones at the southern arms of Sulawesi where these is a mismatch between areas of endemism. More samples will be needed to confirm that the three contact zones running across the eastern arm of Sulawesi, at the 'neck' just North of Central Sulawesi, and Gorontalo Depression (Fig. 4.3) also form effective barriers for *Tylomelania*. Given that the timing of diversification is not fundamentally different in *Tylomelania*, other factors must play a role in causing any mismatch. *Tylomelania* is a strict freshwater dweller with apparent altitudinal distribution limits (see Introduction). This seems to be reflected in the distribution boundaries of the major clades, which are frequently formed by mountain ranges (Fig. 4.3). The contact zones between clade 3 (Fig. 4.2 and 4.3, blue) and clades 1 and 4 (Fig. 4.2 and 4.3, red and green) do fit this hypothesis, as both contact zones are situated in areas where the topography is conducive to the – presumably secondary – contact of these clades. As a caveat, this interpretation is based on present-day topography. However, the orogeny of these mountain barriers most likely set in during the (early) Pliocene (Hall, 2012b) and the bisection of formerly continuous drainages might have occurred practically instantly on a geological time scale.

A serious objection to a strict vicariant hypothesis correlated to orogeny arises from the obvious gaps in our data through insufficient sampling in some regions of Sulawesi, most notably in West Central Sulawesi and along the Eastern arm, but to a lesser degree also in Southeast Sulawesi and along the 'neck' of the island (*Tylomelania* does not occur in North Sulawesi East of the Gorontalo depression). The disjunct distribution of clade 3 (Fig. 4.2 and 4.3, blue), e.g., is most likely a sampling artefact. At the same time and irrespective of sampling issues, the seemingly odd distribution of clade 3 around the Gulf of Bone could have an easy explanation through the recurrent Pleistocene sea-level drops, which increased the land area in the Gulf of Bone considerably (Sathiamurthy & Voris, 2006). To date, *Tylomelania* is the only freshwater taxon from Sulawesi with an island-wide sampling. It will be interesting to see whether our results are corroborated by ongoing studies on other freshwater taxa, such as several groups of gastropods, fishes, atyid shrimps and geocarcinid crabs (M. Glaubrecht, F. Herder, K. von Rintelen & C. Schubart, pers. comm.).

Conclusions

Our molecular divergence time estimates are compatible with the tectonic framework for Sulawesi and thus provide further support for an 'out-of-Australia' vicariance scenario proposed for *Tylomelania* (Glaubrecht & von Rintelen, 2003; von Rintelen & Glaubrecht, 2005; Stelbrink et al., 2012). We also tentatively suggest that the ancestor of *Tylomelania* may have colonized the island from the Sula Spur region of Banggai-Sula, when its western part was fused to the rest of Sulawesi at the Miocene-Pliocene transition. The spatial distribution of genetic diversity as evidenced in *Tylomelania* does at best partially fit the pattern found in several other animal (terrestrial) groups such as macaques, toads or fanged frogs. The primary diversification of *Tylomelania* on Sulawesi into eight major lineages subsequent to the colonization of the island might have been shaped through vicariant events related to the orogeny of the island. Secondary contact between some clades is confined to two restricted areas and has resulted in mitochondrial introgression. These hypotheses could be tested by the future addition of nuclear markers and the sampling of crucial intermediate regions between the known distribution areas of the major lineages of *Tylomelania*.

General Discussion and Conclusions

The previous chapters have clearly demonstrated that time-calibrated trees and palaeontological data (if available) should go hand in hand with information on Southeast Asia's geological and climatic history when studying the biogeographic history of a certain taxonomic group. Furthermore, and more importantly, multi-disciplinary and meta-analytical approaches (among a variety of floral and faunal groups) appear to be the most appropriate strategy for analyzing different levels of biodiversity and more general patterns of colonization, intra-island expansion and diversification across one or more biogeographic areas on the one hand, and for identifying their underlying geological, climatic and biological processes on the other. Consequently, further hypotheses could be tested using newly acquired data from different research fields, taxonomic groups and taxonomically poorly sampled areas. This is a general trend in recent biogeographic studies and has been adopted to other biogeographically interesting areas such as New Zealand (Goldberg et al., 2008; Wallis & Trewick, 2009), Madagascar (Vences et al., 2003; Yoder & Nowak, 2006; Warren et al., 2010), and South America (Hoorn et al., 2010; Ribas et al., 2012; Turchetto-Zolet et al., 2013).

However, the suite of tectonic, palaeogeographic, climatic and phylogenetic reconstructions comes along with a series of uncertainties despite the improvements made in the last decades. Recently published palaeogeographic maps, including high-resolution stratigraphic and oxygen isotope data to infer precise information on past sea-level stands (e.g., Zachos et al., 2001, 2008; Lisiecki & Raymo, 2005; Miller et al., 2005), reveal a detailed picture of an island's or archipelago's geological history through time – compared to the mere identification of tectonic provinces – and thus are more relevant to biogeography-related questions (cf. e.g., Moss & Wilson, 1998; Wilson & Moss, 1999 with e.g., Hall, 2009b, 2012b for details on Sulawesi). Those palaeogeographic maps of complex archipelagos are most likely imprecise on a certain regional or local scale (see e.g., Hall, 2009b) in comparison to insular systems with serial developments of well-dated oceanic islands such as the Hawaiian Archipelago (see General Introduction). However, the combination of both phylogenetic (time-calibrated trees) and palaeogeographic information provides a backbone to test for particular biogeographic hypotheses.

General discussion on previous chapters

The opening review section (Chapter 1) revealed that geological (tectonic movements since the Paleocene) and climatic events (particularly the onset of sea-level fluctuations in the Miocene) had a substantial impact on the archipelago's geological and palaeogeographic configuration and consequently shaped Southeast Asia's biodiversity. These two factors played a vital role for the division of initial biotas (north and south of the Thai-Malay Peninsula) and later triggered

both the accumulation (via immigration and subsequent *in situ* diversification) and emigration of floral and faunal elements. This is most evident in long-term emergent landmasses such as Indochina and Borneo showing high levels of biodiversity in terms of numbers of lineages/species and *in situ* diversifications through time. The rise of the sea level during the middle Miocene climate maximum resulted in extensive inundations of even considerably large islands such as Sumatra and Java. In contrast, Indochina and Borneo have been mostly emergent during this period of time and thus potentially acted as refugia. Molecular clock analyses and ancestral area estimates suggest that both regions have been the main drivers for colonizations across the IAA particularly in the Miocene and Plio-Pleistocene when previously submerged islands became emergent in periods of sea-level drops. Low sea-level stands also had a significant impact on the biotic evolution in one of Southeast Asia's biggest biogeographic regions, namely Sundaland. Widespread landbridges (see Voris, 2000 for Pleistocene reconstructions) are assumed to have facilitated range expansions in various taxa and even gave rise to freshwater fish radiations inhabiting palaeo-drainages now separated by water straits (de Bruyn et al., 2013a; see also Dias et al., 2014 for a global overview on palaeo-drainages). High-resolution vegetational data suggest a 'savannah corridor' for the late Pleistocene running from Indochina south to Borneo and Java (Heaney, 1991). This corridor might not only have caused further range expansions across Southeast Asia as proposed for modern humans, it might have also divided ancestral populations of rainforest-inhabiting species now occurring in the Thai-Malay Peninsula and Sumatra (Bird et al., 2005; but see Slik et al., 2011 for the influence of soil composition in probably forming dispersal barriers).

The interplay of geology and climate-mediated sea-level rises is of peculiar interest in the remaining chapters considering time and origin of Sulawesi's fauna. The meta-analysis performed in Chapter 2 revealed a predominantly Asian origin of Sulawesi groups and again highlights the need for palaeogeographic information when distribution patterns are studied. Two major geological events were identified representing constraints on vicariant processes and gave rise to present-day Sulawesi: 1) the opening of the Makassar Strait c. 45 Mya separating West Sulawesi from Sundaland (Borneo; Sunda Shelf), and 2) the extension and separation of the Sula Spur from the Australian margin (Sahul Shelf) starting c. 15 Mya. As a result, Sulawesi's biota (assuming terrestrial and freshwater refugia on the Sula Spur) could show direct affinities to each of the two regions (shelves) being directly derived from its ancestral population in the course of the respective vicariant process. However, testing such hypotheses necessarily requires phylogenetic information on Sulawesi taxa and their relatives including estimated divergence times. Only few datasets showed divergence time estimates predating or coinciding with age estimates of the respective geological event – congruent with a vicariant origin (see below) – while in the majority of datasets analyzed the respective splits clearly postdated these two events and thus refuted such a scenario for those taxonomic groups. Here,

several terrestrial and freshwater taxa might have colonized Sulawesi by dispersal when low sea-level stands in the Plio-Pleistocene timeframe facilitated intra-archipelago range expansion. All these factors explain why it might be more reasonable to propose just two demarcation lines separating Asian and Australian faunas, namely Wallace's Line and Lydekker's Line, representing the western and eastern borders of Wallacea and thus the region of non-continental islands. Moreover, due to its previous physical contact to both shelves and the resulting exchange of floral and faunal elements, this biogeographic area thus might indeed be considered a 'transition zone'. However, the comparatively long geographic and hence genetic isolation from surrounding islands and shelves gave rise to a peculiar flora and fauna quite distinct from Asia's and Australia's biota. Recent biogeographic studies show that Sulawesi (and Wallacea in general) have been mainly colonized from Indochina, the Philippines, and the Australian margin, while interestingly, colonizations from nearby Borneo are comparatively infrequent (see colonization routes in Chapter 1, Fig. 1.4). Hence, for taxa with low dispersal capacities (i.e., non-volant terrestrial and freshwater-restricted taxa), particularly Wallace's Line represents an important dispersal barrier likely due to the strong ocean current running southwards along the Makassar Strait.

In Chapter 3 and 4, two freshwater groups were studied in detail, namely pachychilid snails and telmatherinid fishes endemic to Sulawesi and both have experienced a remarkable intra-lacustrine radiation inhabiting the Malili Lakes system (and Lake Poso). For both groups, Australian (New Guinean) affinities have been previously suggested based on molecular phylogenetic analyses and anatomical features (snails: Glaubrecht & von Rintelen, 2003; von Rintelen & Glaubrecht, 2005; fishes: Sparks & Smith, 2004; Bloom et al., 2012). However, the present-day distribution of close relatives and a life history without any developmental stages tolerant of brackish or saline waters raised the question how ancestral populations of these two groups could colonize Sulawesi. The proposed sister group of the snail genus *Tylomelania* is only represented by two single species endemic to two Torres Strait islands (Prince of Wales Island, Hammond Island). Different phylogenetic hypotheses were proposed for atherinomorph fishes, however, these studies did not include any or all Sulawesi representatives and thus sister group relationships remained unclear so far. Palaeogeographic reconstructions and time-calibrated phylogenies make it seem plausible that the Sula Spur (the continuation of the Australian continental margin in New Guinea) has acted as a so-called 'terrane raft' for ancestral populations of both taxonomic groups (vicariant scenario). However, this hypothesis needs to be tested against additional time-calibrated phylogenies of other Sulawesi freshwater taxa with Australian affinities; it not only implies continuously emergent areas, but more importantly, the continuous existence of freshwater refugia during the westward movement until the Sula Spur finally collided with West Sulawesi.

Potential pitfalls in biogeographic analyses

The last issue once again underlines the importance of assessing the reliability of several aspects in a given biogeographic framework. One important point with a high potential for errors concerns taxonomic and geographic coverage, particularly in large-scale meta-analyses as performed in Chapter 1. In this study, several datasets have been *a priori* excluded from the molecular meta-analysis due to questionable calibration points (see below), missing potential outgroup taxa, but more importantly, sparse sampling across Southeast Asia. Sampling artefacts may result in incomplete (time-calibrated) phylogenies and thus directly affect the results of ancestral area estimates used to reconstruct colonization routes. As emphasized in Chapter 1, some regions in the IAA are well covered in the molecular datasets, while others such as Palawan and New Guinea are likely under-represented though these two areas show high levels of diversity and have certainly played a crucial role for inter-island colonization and *in situ* diversification. This is related to a limited accessibility to these regions and a poor availability of relevant data, i.e., DNA sequences based standard genetic markers.

A second aspect that has been only briefly discussed in each of the chapters is the problems associated with performing time-calibrated phylogenetic analyses. Despite many improvements in the field of historical biogeography – including species tree reconstructions based on multiple gene genealogies (Liu & Pearl, 2007; Heled & Drummond, 2010), the inference of demographic histories within populations and species flocks (Ho & Shapiro, 2011), the use of genomic data in molecular clock analyses (Ho, 2014), and refined ancestral area estimates (Matzke, 2013) – the choice of calibration points applied represents the most important source of dating errors (see e.g., Donoghue & Benton, 2007; Ho, 2007; Sanders & Lee, 2007; Ho et al., 2008; Ho & Philipps, 2009; Lukoschek et al., 2012; Warnock et al., 2012; Hipsley & Müller, 2014).

One of the most simple calibration strategies is probably the use of so-called secondary (indirect) calibration points. Although this approach only requires the conversion of posterior distributions (divergence time estimates) from source publications into calibration priors for subordinated analyses, the underlying source calibration might be questionable and resulting credibility intervals can be comparatively large. This methodology has been also applied in Chapter 3 to date Sulawesi's telmatherinid radiation; however, inferred divergence times clearly overestimated any age estimates of Sulawesi (and thus its inhabiting freshwater fishes) possibly related to the choice of fossils in the source. Consequently, indirect calibration points might be helpful for initial rough age estimates but source calibration points need to be carefully evaluated.

Another approach is to use geological calibration points. In Chapter 3, the formation of the Central Highlands in New Guinea was used to infer age estimates in telmatherinid fishes. This type of calibration is prone to circular reasoning if a certain geological event used for

calibrating a phylogeny reflects a particular biogeographic scenario to be tested (e.g., vicariant constraint). Furthermore, geological events such as the opening of water straits, the formation of mountain ranges and the closure of an isthmus represent ongoing processes and thus impede the precise dating of geographic and genetic isolation of ancestral populations. Hence, geological calibration points are potentially useful to infer maximum ages for taxonomic groups with low dispersal capacities inhabiting certain (oceanic) islands such as the Hawaiian Archipelago. In contrast, applying geological events for dating geographically remote groups and geographically (and genetic) isolated populations only works to a limited extent.

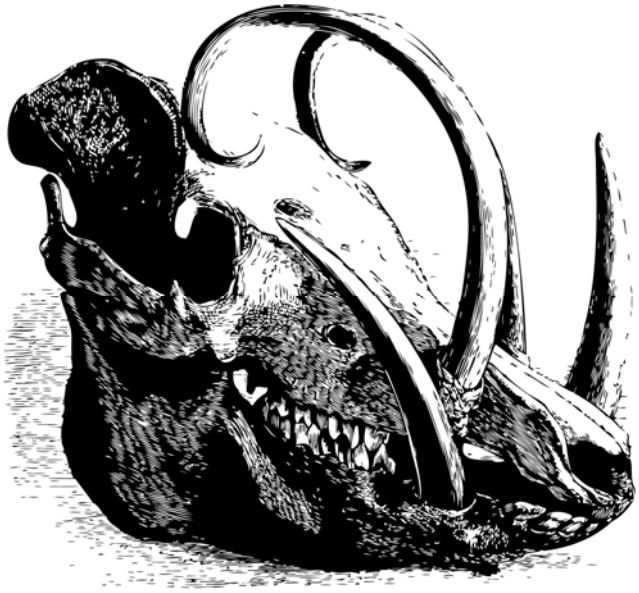
A widely used method is the application of substitution rates as suggested for mitochondrial markers in arthropods (insects) and birds (Brower, 1994; Weir & Schluter, 2008). Substitution rates are rarely based on direct observations but most frequently used if other external calibration points are missing. Most of the datasets analysed in the preceding meta-analyses (particularly Chapter 1) were performed using general rates to infer general diversification and colonization patterns among different taxa (e.g., birds, insects). However, the use of non-specific substitution rates should always be used with caution because this might potentially lead to misinterpretations when testing for a certain biogeographic hypothesis (see e.g., Phillips et al., 2013). One remarkable example, ricefishes of the genus *Oryzias* endemic to Sulawesi, has been highlighted in Chapter 2. While phylogenetic relationships clearly indicated an Asian origin for this group, estimated divergence times proposed by the source authors (29-32 Mya) suggested a vicariant origin for these fishes (though this scenario appears to be unlikely given the palaeogeographic reconstructions). However, the substitution rate applied by the source authors appeared questionable because it was derived from distantly related Antarctic ice-fishes (Notothenioidei). Moreover, cold-water adapted species are characterized by a generally lower metabolism and thus decreased mutation rates, which can result in comparatively old divergence time estimates if applied to tropical taxa with an increased metabolism. Thus, a different substitution rate derived from closer related cyprinodont fishes was used for a re-analysis resulting in a considerably more recent origin (3-7 Mya) and hence rejecting a vicariant scenario. This single example reveals the main difficulty of not using taxon-, marker, and model-specific substitution rates, for which unfortunately only few data is available (but see e.g., Wilke et al., 2009 for freshwater invertebrates, and Burridge et al., 2008 for freshwater and marine fishes).

The final strategy discussed here in greater detail, fossil calibration, appears to be a more robust method compared to the abovementioned approaches and likely becomes more important for biogeographic questions in the near future. Fossil and pollen records shed light on the early evolution of floras and faunas and uncover past changes in vegetation, habitat composition, and climate through time. However, the fossil record in Southeast Asia is heavily biased in terms of geographic distribution, taxonomic coverage and geological age (see also

Chapter 1). While the Mesozoic fossil record shows some plant and (marine) vertebrate taxa, several fossils are found in the Eocene and Oligocene mainly represented by mammals such as dermopterans, rodents and primates. However, these fossil remains are mainly located in mainland Southeast Asia (Thailand, Laos and Myanmar) and some few, but famous, cave sites like the Niah and Madai caves in Borneo primarily containing hominins and birds from the Pleistocene and Holocene (e.g., van den Bergh et al., 2009; Louys & Meijaard, 2010; Meijer, 2014). This patchy distribution across Southeast Asia is probably related to the complex tectonic movements and pronounced sea-level fluctuations resulting in generally instable environments and long-term inundations of low-elevation areas, for example, found in Java and Sumatra (cf. Chapter 1 with Fig. 1 in Louys & Meijaard, 2010 and Meijer, 2014). Some of these fossilized specimens are potentially useful to infer divergence times within a given phylogenetic tree by dating one or more nodes if age determination and phylogenetic assignment are carefully evaluated. However, only few fossils are yet recorded for invertebrates, larger mammals and freshwater taxa examined in the above meta-analyses. This fact not only requires the use of other (potentially less robust) calibration strategies such as non-specific substitution rates, it also prevents testing for widespread extinction in different taxonomic groups for which a vicariant scenario has been previously proposed (e.g., pachychilid snails and telmatherinid fishes).

Outlook

Southeast Asia represents a challenging but also extremely interesting region to study historical biogeography due to its complex geological and biological nature. The present thesis has shown that future research should focus on both large-scale meta-analyses to infer general patterns among a variety of groups but also on small-scale studies investigating local or regional diversification patterns and demographic histories in species flocks and populations. Moreover, these studies can contribute to the understanding of biogeographic borders and the factors and processes that have triggered colonizations across such demarcation lines. However, this requires the collection of more data (taxonomic and geographic coverage), particularly including organisms with low dispersal capacities (such as freshwater taxa), which are not well represented in current studies and species richness maps compared to birds and mammals. Further research is needed to investigate the influence of orogenetic events, palaeo-drainages and past refugia (i.e., emergent areas during high sea-level stands) on the population structure of Sulawesi's endemic species by using refined palaeogeographic data and additional time-calibrated phylogenies. Future drilling projects planned for the Malili lakes and Lake Poso will yield more precise age estimates and thus will provide insights into the lake's history and possibly even into its fossil fauna. More importantly, such data will allow identifying and dating congruent patterns of diversification and colonization among the various lacustrine radiations and will also help to relate these patterns to past geological processes.



The above picture shows the skull of a babirusa, modified from: Wallace A.R. (1869) *The Malay Archipelago, the land of the orang-utan and the bird of paradise a narrative of travel with studies of man and nature*. Macmillan, London (p. 283).

References

Aarn W.I., Ivantsoff W. & Kottelat M. (1998) Phylogenetic analysis of Telmatherinidae (Teleostei: Atherinomorpha), with description of *Marosatherina*, a new genus from Sulawesi. *Ichthyological Exploration of Freshwaters*, **9**, 311–323.

Abell R., Thieme M.L., Revenga C., Bryer M., Kottelat M., Bogutskaya N., Coad B., Mandrak N., Contreras Balderas S., Bussing W., Stiassny M.L.J., Skelton P., Allen G.R., Unmack P., Naseka A., Ng R., Sindorf N., Robertson J., Armijo E., Higgins J. V, Heibel T.J., Wikramanayake E., Olson D., Lopéz H.L., Reis R.E., Lundberg J.G., Sabaj Pérez M.H. & Petry P. (2008) Freshwater ecoregions of the world: a new map of biogeographic units for freshwater biodiversity conservation. *BioScience*, **58**, 403–414.

Achard F., Eva H.D., Stibig H.-J., Mayaux P., Gallego J., Richards T. & Malingreau J.-P. (2002) Determination of deforestation rates of the world's humid tropical forests. *Science*, **297**, 999–1002.

Aldrich J.B., Rinehart G.P., Ridwan S. & Schuepbach M.A. (1996) Paleogene basin architecture of the Sunda and Asri Basins and associated non-marine sequence stratigraphy. *Proceedings of the International Symposium on Sequence Stratigraphy in S.E. Asia*, 261–287.

Alfaro M.E., Karns D.R., Voris H.K., Brock C.D. & Stuart B.L. (2008) Phylogeny, evolutionary history, and biogeography of Oriental-Australian rear-fanged water snakes (Colubroidea: Homalopsidae) inferred from mitochondrial and nuclear DNA sequences. *Molecular Phylogenetics and Evolution*, **46**, 576–593.

Arbogast B.S. & Kenagy G.J. (2001) Comparative phylogeography as an integrative approach to historical biogeography. *Journal of Biogeography*, **28**, 819–825.

Armstrong R.A. (2014) When to use the Bonferroni correction. *Ophthalmic & Physiological Optics*, in press.

Avise J.C. (2004) *Molecular markers, natural history and evolution.* Sinauer, Sunderland, Mass.

Ayala F.J. (1999) Molecular clock mirages. *BioEssays*, **21**, 71–75.

Baker A.J., Huynen L.J., Haddrath O., Millar C.D. & Lambert D.M. (2005) Reconstructing the tempo and mode of evolution in an extinct clade of birds with ancient DNA: the giant moas of New Zealand. *Proceedings of the National Academy of Sciences USA*, **105**, 8257–8262.

Balke M., Ribera I., Hendrich L., Miller M.A., Sagata K., Posman A., Vogler A.P. & Meier R. (2009) New Guinea highland origin of a widespread arthropod supertramp. *Proceedings of the Royal Society of London B*, **276**, 2359–2367.

Balke M., Ribera I. & Vogler A.P. (2004) MtDNA phylogeny and biogeography of Copelatinae, a highly diverse group of tropical diving beetles (Dytiscidae). *Molecular Phylogenetics and Evolution*, **32**, 866–880.

Bargelloni L., Marcato S., Zane L. & Patarnello T. (2000) Mitochondrial phylogeny of notothenioids: a molecular approach to Antarctic fish evolution and biogeography. *Systematic Biology*, **49**, 114–129.

Bartish I.V., Antonelli A., Richardson J.E. & Swenson U. (2011) Vicariance or long-distance dispersal: historical biogeography of the pantropical subfamily Chrysophylloideae (Sapotaceae). *Journal of Biogeography*, **38**, 177–190.

Beard K.C., Marivaux L., Chaimanee Y., Jaeger J.-J., Marandat B., Tafforeau P., Soe A.N., Tun S.T. & Kyaw A.A. (2009) A new primate from the Eocene Pondaung Formation of Myanmar and the monophyly of Burmese amphipithecids. *Proceedings of the Royal Society of London B*, **276**, 3285–3294.

Benton M.J. & Donoghue P.C.J. (2007) Paleontological evidence to date the tree of life. *Molecular Biology and Evolution*, **24**, 26–53.

Bess E.C., Catanach T.A. & Johnson K.P. (2014) The importance of molecular dating analyses for inferring Hawaiian biogeographical history: a case study with bark lice (Psocidae: *Ptycta*). *Journal of Biogeography*, **1**, 158–167.

Bird M.I., Taylor D. & Hunt C. (2005) Palaeoenvironments of insular Southeast Asia during the Last Glacial Period: a savanna corridor in Sundaland? *Quaternary Science Reviews*, **24**, 2228–2242.

Birks S.M. & Edwards S.V. (2002) A phylogeny of the megapodes (Aves: Megapodiidae) based on nuclear and mitochondrial DNA sequences. *Molecular Phylogenetics and Evolution*, **23**, 408–421.

Blackburn D.C., Bickford D.P., Diesmos A.C., Iskandar D.T. & Brown R.M. (2010) An ancient origin for the enigmatic flat-headed frogs (Bombinatoridae: *Barbourula*) from the islands of Southeast Asia. *PLoS ONE*, **5**, e12090.

Bloom D.D., Unmack P.J., Gosztonyi A.E., Piller K.R. & Lovejoy N.R. (2012) It's a family matter: molecular phylogenetics of Atheriniformes and the polyphyly of the surf silversides (Family: Notocheiridae). *Molecular Phylogenetics and Evolution*, **62**, 1025–1030.

Bridle J.R., Garn A.K., Monk K.A. & Butlin R.K. (2001) Speciation in *Chitaura* grasshoppers (Acrididae: Oxyinae) on the island of Sulawesi: colour patterns, morphology and contact zones. *Biological Journal of the Linnean Society*, **72**, 373–390.

Bridle J.R., Pedro P.M. & Butlin R.K. (2004) Habitat fragmentation and biodiversity: testing for the evolutionary effects of refugia. *Evolution*, **58**, 1394–1396.

Bromham L. & Penny D. (2003) The modern molecular clock. *Nature Reviews Genetics*, **4**, 216–224.

Brower A.V.Z. (1994) Rapid morphological radiation and convergence among races of the butterfly *Heliconius erato* inferred from patterns of mitochondrial DNA evolution. *Proceedings of the National Academy of Sciences USA*, **91**, 6491–6495.

Brown P., Sutikna T., Morwood M.J., Soejono R.P., Jatmiko, Saptomo E.W. & Due R.A. (2004) A new small-bodied hominin from the Late Pleistocene of Flores, Indonesia. *Nature*, **431**, 1055–1061.

Brown R.M., Linkem C.W., Siler C.D., Sukumaran J., Esselstyn J.A., Diesmos A.C., Iskandar D.T., Bickford D., Evans B.J., McGuire J.A., Grismer L., Supriatna J. & Andayani N. (2010) Phylogeography and historical demography of *Polypedates leucomystax* in the islands of Indonesia and the Philippines: Evidence for recent human-mediated range expansion? *Molecular Phylogenetics and Evolution*, **57**, 598–619.

Brown R.M., Siler C.D., Oliveros C.H., Esselstyn J.A., Diesmos A.C., Hosner P.A., Linkem C.W., Barley A.J., Oaks J.R., Sanguila M.B., Welton L.J., Blackburn D.C., Moyle R.G., Townsend Peterson A. & Alcala A.C. (2013) Evolutionary processes of diversification in a model island archipelago. *Annual Review of Ecology, Evolution, and Systematics*, **44**, 411–435.

Bunce M., Szulkin M., Lerner H.R.L., Barnes I., Shapiro B., Cooper A. & Holdaway R.N. (2005) Ancient DNA provides new insights into the evolutionary history of New Zealand's extinct giant eagle. *PLoS Biology*, **3**, e9.

Burrett C., Duhig N., Berry R. & Varne R. (1991) Asian and South-western Pacific continental terranes derived from Gondwana, and their biogeographic significance. *Australian Systematic Botany*, **4**, 13–24.

Burridge C.P., Craw D., Fletcher D. & Waters J.M. (2008) Geological dates and molecular rates: fish DNA sheds light on time dependency. *Molecular Biology and Evolution*, **25**, 624–633.

Butlin R.K., Walton C., Monk K.A., Bridle J.R., Hall R. & Holloway J.D. (1998) Biogeography of Sulawesi grasshoppers, genus *Chitaura*, using DNA sequence data. *Biogeography and Geological Evolution of SE Asia* (ed. by R. Hall and D. Holloway), pp. 355–359. Backhuys Publishers, Leiden.

Cannon C.H., Morley R.J. & Bush A.B.G. (2009) The current refugial rainforests of Sundaland are unrepresentative of their biogeographic past and highly vulnerable to disturbance. *Proceedings of the National Academy of Sciences USA*, **106**, 11188–11193.

Catullo G., Masi M., Falcucci A., Maiorano L., Rondinini C. & Boitani L. (2008) A gap analysis of Southeast Asian mammals based on habitat suitability models. *Biological Conservation*, **141**, 2730–2744.

Cerwenka A.F., Wedekind J.D., Hadiaty R.K., Schliewen U.K. & Herder F. (2012) Alternative egg-feeding tactics in *Telmatherina sarasinorum*, a trophic specialist of Lake Matano's evolving sailfin silversides fish radiation. *Hydrobiologia*, **693**, 131–139.

Clement M., Posada D. & Crandall K.A. (2000) TCS: a computer program to estimate gene genealogies. *Molecular Ecology*, **9**, 1657–1659.

Clouse R.M. & Giribet G. (2010) When Thailand was an island – the phylogeny and biogeography of mite harvestmen (Opiliones, Cyphophthalmi, Stylocellidae) in Southeast Asia. *Journal of Biogeography*, **37**, 1114–1130.

Cohen A.S. & Johnston M.R. (1987) Speciation in brooding and poorly dispersing lacustrine organisms. *Palaios*, **2**, 426–435.

Corlett R.T. (2007) The impact of hunting on the mammalian fauna of tropical Asian forests. *Biotropica*, **39**, 292–303.

Cowie R.H. & Holland B.R. (2006) Dispersal is fundamental to biogeography and the evolution of biodiversity on oceanic islands. *Journal of Biogeography*, **33**, 193–198.

Cowie R.H. & Holland B.S. (2008) Molecular biogeography and diversification of the endemic terrestrial fauna of the Hawaiian Islands. *Philosophical Transactions of the Royal Society of London B*, **363**, 3363–3376.

Cox C.B. (2001) The biogeographic regions reconsidered. *Journal of Biogeography*, **28**, 511–523.

Cranbrook, Earl of (2010) Late quaternary turnover of mammals in Borneo: the zooarchaeological record. *Biodiversity and Conservation*, **19**, 373–391.

Crawford A.J. (2003) Relative rates of nucleotide substitution in frogs. *Journal of Molecular Evolution*, **57**, 636–641.

Crisp M.D., Trewick S.A. & Cook L.G. (2011) Hypothesis testing in biogeography. *Trends in Ecology & Evolution*, **26**, 66–72.

Croizat L., Nelson G. & Rosen D.E. (1974) Centres of origin and related concepts. *Systematic Zoology*, **23**, 265–287.

Crottini A., Madsen O., Poux C., Strauß A., Vieites D.R. & Vences M. (2012) Vertebrate time-tree elucidates the biogeographic pattern of a major biotic change around the K-T boundary in Madagascar. *Proceedings of the National Academy of Sciences USA*, **109**, 5358–5363.

de Beaufort L.F. (1926) *Zoögeographie van den Indischen Archipel.* Bohn, Haarlem.

de Boer A.J. & Duffels J.P. (1996) Historical biogeography of the cicadas of Wallacea, New Guinea and the West Pacific: a geotectonic explanation. *Palaeogeography, Palaeoclimatology, Palaeoecology*, **124**, 153–177.

de Bruyn M., Hoelzel A.R., Carvalho G.R. & Hofreiter M. (2011) Faunal histories from Holocene ancient DNA. *Trends in Ecology & Evolution*, **26**, 405–413.

de Bruyn M., Rüber L., Nylinder S., Stelbrink B., Lovejoy N.R., Lavoué S., Tan H.H., Nugroho E., Wowor D., Ng P.K.L., Siti Azizah M.N., von Rintelen T., Hall R. & Carvalho G.R. (2013a) Paleo-drainage basin connectivity predicts evolutionary relationships across three Southeast Asian biodiversity hotspots. *Systematic Biology*, **62**, 398–410.

de Bruyn M., Stelbrink B., Page T.J., Phillips M.J., Lohman D.J., Albrecht C., Hall R., von Rintelen K., Ng P.K.L., Shih H.-T., Carvalho G.R. & von Rintelen T. (2013b) Time and space in biogeography: response to Parenti & Ebach (2013). *Journal of Biogeography*, **40**, 2204–2206.

de Bruyn M., von Rintelen T., von Rintelen K., Mather P.B. & Carvalho G.R. (2012) Molecular biogeography and phylogeography of the freshwater fauna of the Indo-Australian Archipelago. *Biotic evolution and environmental change in Southeast Asia* (ed. by D.J. Gower, K.G. Johnson, J.E. Richardson, B.R. Rosen, L. Rüber, and S.T. Williams), pp. 316–346. Cambridge University Press, Cambridge.

de Graciansky P.-C., Hardenbol J., Jacquin T. & Vail P.R. (1998) *Mesozoic and Cenozoic sequence stratigraphy of European basins.* Society of Economic Paleontologists and Mineralogists.

de Queiroz A. (2005) The resurrection of oceanic dispersal in historical biogeography. *Trends in Ecology & Evolution*, **20**, 68–73.

Dekker R.W.R.J. (2007) Distribution and speciation of megapodes (Megapodiidae) and subsequent development of their breeding behaviour. *Biogeography, Time, and Place: Distributions, Barriers, and Islands* (ed. by W. Renema), pp. 93–102. Springer, Dordrecht.

Dias M.S., Oberdorff T., Hugueny B., Leprieur F., Jézéquel C., Cornu J.-F., Brosse S., Grenouillet G. & Tedesco P.A. (2014) Global imprint of historical connectivity on freshwater fish biodiversity. *Ecology Letters*, **17**, 1130–1140.

Dickerson R.E. (1928) *Distribution of Life in the Philippines.* Bureau of Sciences, Manila.

Donoghue M.J. & Moore B.R. (2003) Towards an integrative historical biogeography. *Integrative and Comparative Biology*, **43**, 261–270.

Donoghue P.C.J. & Benton M.J. (2007) Rocks and clocks: calibrating the Tree of Life using fossils and molecules. *Trends in Ecology & Evolution*, **22**, 424–431.

Drummond A.J., Ho S.Y.W., Phillips M.J. & Rambaut A. (2006) Relaxed phylogenetics and dating with confidence. *PLoS Biology*, **4**, e88.

Drummond A.J. & Rambaut A. (2007) BEAST: Bayesian evolutionary analysis by sampling trees. *BMC Evolutionary Biology*, **7**, 214.

Drummond A.J., Rambaut A., Shapiro B. & Pybus O.G. (2005) Bayesian coalescent inference of past population dynamics from molecular sequences. *Molecular Biology and Evolution*, **22**, 1185–1192.

Drummond A.J. & Suchard M.A. (2010) Bayesian random local clocks, or one rate to rule them all. *BMC Biology*, **8**, 114.

Drummond A.J., Suchard M.A., Xie D. & Rambaut A. (2012) Bayesian phylogenetics with BEAUti and the BEAST 1.7. *Molecular Biology and Evolution*, **29**, 1969–1973.

Ducrocq S. (1996) The Eocene terrestrial mammal from Timor, Indonesia. *Geological Magazine*, **133**, 763–766.

Ebach M.C. & Morrone J.J. (2005) Forum on historical biogeography: what is cladistic biogeography? *Journal of Biogeography*, **32**, 2179–2183.

Echelle A.A., Carson E.W., Echelle A.F., Van Den Bussche R.A., Dowling T.E. & Meyer A. (2005) Historical biogeography of the New-World pupfish genus *Cyprinodon* (Teleostei: Cyprinodontidae). *Copeia*, **2005**, 320–339.

Edgar R.C. (2004) MUSCLE: a multiple sequence alignment method with reduced time and space complexity. *BMC Bioinformatics*, **5**, 113.

Esselstyn J.A. & Oliveros C.H. (2010) Colonization of the Philippines from Taiwan: a multi-locus test of the biogeographic and phylogenetic relationships of isolated populations of shrews. *Journal of Biogeography*, **37**, 1504–1514.

Esselstyn J.A., Timm R.M. & Brown R.M. (2009) Do geological or climatic processes drive speciation in dynamic archipelagos? The tempo and mode of diversification in Southeast Asian shrews. *Evolution*, **63**, 2595–2610.

Evans B.J., Brown R.M., McGuire J.A., Supriatna J., Andayani N., Diesmos A., Iskandar D., Melnick D.J. & Cannatella D.C. (2003a) Phylogenetics of fanged frogs: testing biogeographical hypotheses at the interface of the Asian and Australian faunal zones. *Systematic Biology*, **52**, 794–819.

Evans B.J., McGuire J.A., Brown R.M., Andayani N. & Supriatna J. (2008) A coalescent framework for comparing alternative models of population structure with genetic data: evolution of Celebes toads. *Biology Letters*, **4**, 430–433.

Evans B.J., Morales J.C., Supriatna J. & Melnick D.J. (1999) Origin of the Sulawesi macaques (Cercopithecidae, *Macaca*) as inferred from a mitochondrial DNA phylogeny. *Biological Journal of the Linnean Society*, **66**, 539–560.

Evans B.J., Pin L., Melnick D.J. & Wright S.I. (2010) Sex-linked inheritance in macaque monkeys: implications for effective population size and dispersal to Sulawesi. *Genetics*, **185**, 923–937.

Evans B.J., Supriatna J., Andayani N. & Melnick D.J. (2003b) Diversification of Sulawesi macaque monkeys: decoupled evolution of mitochondrial and autosomal DNA. *Evolution*, **57**, 1931–1946.

Evans B.J., Supriatna J., Andayani N., Setiadi M.I., Cannatella D.C. & Melnick D.J. (2003c) Monkeys and toads define areas of endemism on Sulawesi. *Evolution*, **57**, 1436–1443.

Folinsbee K.E. & Evans D.C. (2012) A protocol for temporal calibration of general area cladograms. *Journal of Biogeography*, **39**, 688–697.

Friedman M., Keck B.P., Dornburg A., Eytan R.I., Martin C.H., Darrin C., Wainwright P.C., Near T.J. & Hulsey C.D. (2013) Molecular and fossil evidence place the origin of cichlid fishes long after Gondwanan rifting. *Proceedings of the Royal Society of London B*, **280**, 1–8.

Funk D.J. & Omland K.E. (2003) Species-level paraphyly and polyphyly: frequency, causes, and consequences, with insights from animal mitochondrial DNA. *Annual Review of Ecology and Systematics*, **34**, 397–423.

Garrard R.A., Supandjono J.B. & Surono (1988) The geology of the Banggai-Sula microcontinent, eastern Indonesia. *Proceedings Indonesian Petroleum Association*, **17th Annual**, 23–52.

Gilbert M.T.P., Drautz D.I., Lesk A.M., Ho S.Y.W., Qi J., Ratan A., Hsu C.-H., Sher A., Dalén L., Götherström A., Tomsho L.P., Rendulic S., Packard M., Campos P.F., Kuznetsova T.V., Shidlovskiy F., Tikhonov A., Willerslev E., Iacumin P., Buigues B., Ericson P.G.P., Germonpré M., Kosintsev P., Nikolaev V., Nowak-Kemp M., Knight J.R., Irzyk G.P., Perbost C.S., Fredrikson K.M., Harkins T.T., Sheridan S., Miller W. & Schuster S.C. (2008) Intraspecific phylogenetic analysis of Siberian woolly mammoths using complete mitochondrial genomes. *Proceedings of the National Academy of Sciences USA*, **105**, 8327–8332.

Gillespie R.G., Baldwin B.G., Waters J.M., Fraser C.I., Nikula R. & Roderick G.K. (2012) Long-distance dispersal: a framework for hypothesis testing. *Trends in Ecology & Evolution*, **27**, 47–56.

Giribet G. & Boyer S.L. (2010) "Moa"s Ark' or "Goodbye Gondwana": is the origin of New Zealand's terrestrial invertebrate fauna ancient, recent, or both? *Invertebrate Systematics*, **24**, 1–8.

Glaubrecht M. (1996) *Evolutionsökologie und Systematik am Beispiel von Süß- und Brackwasserschnecken (Mollusca: Caenogastropoda: Cerithioidea): Ontogenese-Strategien, paläontologische Befunde und historische Zoogeographie.* Backhuys, Leiden.

Glaubrecht M. & von Rintelen T. (2003) Systematics, molecular genetics and historical zoogeography of the viviparous freshwater gastropod *Pseudopotamis* (Cerithioidea, Pachychilidae): a relic on the Torres Strait Islands, Australia. *Zoologica Scripta*, **32**, 415–435.

Glaubrecht M. & von Rintelen T. (2008) The species flocks of lacustrine gastropods: *Tylomelania* on Sulawesi as models in speciation and adaptive radiation. *Hydrobiologia*, **615**, 181–199.

Goldberg E.E., Lancaster L.T. & Ree R.H. (2011) Phylogenetic inference of reciprocal effects between geographic range evolution and diversification. *Systematic Biology*, **60**, 451–465.

Goldberg J., Trewick S.A. & Paterson A.M. (2008) Evolution of New Zealand's terrestrial fauna: a review of molecular evidence. *Philosophical Transactions of the Royal Society of London B*, **363**, 3319–3334.

Gorog A.J., Sinaga M.H. & Engstrom M.D. (2004) Vicariance or dispersal? Historical biogeography of three Sunda shelf murine rodents (*Maxomys surifer*, *Leopoldamys sabanus* and *Maxomys whiteheadi*). *Biological Journal of the Linnean Society*, **81**, 91–109.

Gower D., Johnson K.G., Rosen B.R., Richardson J., Rüber L. & Williams S.T. (2012) *Biotic evolution and environmental change in Southeast Asia.* Cambridge University Press, Cambridge.

Grandcolas P., Murienne J., Robillard T., Desutter-Grandcolas L., Jourdan H., Guilbert E. & Deharveng L. (2008) New Caledonia: a very old Darwinian island? *Philosophical Transactions of the Royal Society of London B*, **363**, 3309–3317.

Graur D. & Martin W. (2004) Reading the entrails of chickens: molecular timescales of evolution and the illusion of precision. *Trends in Genetics*, **20**, 80–86.

Gray S.M., Dill L.M., Tantu F.Y., Loew E.R., Herder F. & McKinnon J.S. (2008) Environment-contingent sexual selection in a colour polymorphic fish. *Proceedings of the Royal Society of London B*, **275**, 1785–1791.

Grummer J.A., Bryson Jr. R.W. & Reeder T.W. (2014) Species delimitation using Bayes factors: simulations and application to the Sceloporus scalaris species group (Squamata: Phrynosomatidae). *Systematic Biology*, **63**, 119–133.

Hall R. (1996) Reconstructing Cenozoic SE Asia. *Tectonic evolution of Southeast Asia* (ed. by R. Hall and D.J. Blundell), pp. 153–184. Geological Society, London.

Hall R. (1997) Cenozoic plate tectonic reconstructions of SE Asia. *Petroleum Geology of Southeast Asia* (ed. by A.J. Fraser, S.J. Matthews, and R.W. Murphy), pp. 11–23. Geological Society, London.

Hall R. (2001) Cenozoic reconstructions of SE Asia and the SW Pacific: changing patterns of land and sea. *Faunal and floral migrations and evolution in SE Asia-Australasia* (ed. by I. Metcalfe, J.M.B. Smith, M. Morwood, and I. Davidson), pp. 35–56. Balkema, Lisse.

Hall R. (2002) Cenozoic geological and plate tectonic evolution of SE Asia and the SW Pacific: computer-based reconstructions, model and animations. *Journal of Asian Earth Sciences*, **20**, 353–431.

Hall R. (2009a) Continental growth at the Indonesian margins of southeast Asia. *Arizona Geological Society Digest*, **22**, 245–258.

Hall R. (2009b) Southeast Asia's changing palaeogeography. *Blumea*, **54**, 148–161.

Hall R. (2011) Australia-SE Asia collision: plate tectonics and crustal flow. *The Southeast Asian Gateway: history and tectonics of Australia-Asia collision* (ed. by R. Hall, M.A. Cottam, and M.E.J. Wilson), pp. 75–109. The Geological Society of London, London.

Hall R. (2012a) Late Jurassic-Cenozoic reconstructions of the Indonesian region and the Indian Ocean. *Tectonophysics*, **570-571**, 1–41.

Hall R. (2012b) Sundaland and Wallacea: geology, plate tectonics and palaeogeography. *Biotic evolution and environmental change in Southeast Asia* (ed. by D.J. Gower, K.G. Johnson, J.E. Richardson, B.R. Rosen, L. Rüber, and S.T. Williams), pp. 32–78. Cambridge University Press, Cambridge.

Hall R. (2013) The palaeogeography of Sundaland and Wallacea since the Late Jurassic. *Journal of Limnology*, **72**, 1–17.

Hamilton W. (1979) Tectonics of the Indonesian Region. *U.S. Geological Survey Professional Paper*, **1078**, 1–345.

Hanebuth T.J.J., Voris H.K., Yokoyama Y., Saito Y. & Okuno J. (2011) Formation and fate of sedimentary depocentres on Southeast Asia's Sunda Shelf over the past sea-level cycle and biogeographic implifications. *Earth-Science Reviews*, **104**, 92–110.

Haq B.U., Hardenbol J. & Vail P.R. (1987) Chronology of fluctuating sea levels since the Triassic. *Science*, **235**, 1156–1167.

Harley M.M. & Morley R.J. (1995) Ultrastructural studies of some fossil and extant palm pollen, and the reconstruction of the biogeographical history of subtribes Iguanurinae and Calaminae. *Review of Palaeobotany and Palynology*, **85**, 153–182.

Hasegawa M., Kishino H., Hayasaka K. & Horai S. (1990) Mitochondrial DNA evolution in primates: transition rate has been extremely low in the lemur. *Journal of Molecular Evolution*, **31**, 113–121.

Hassanin A. & Douzery E.J.P. (1999) The tribal radiation of the family Bovidae (Artiodactyla) and the evolution of the mitochondrial cytochrome b gene. *Molecular Phylogenetics and Evolution*, **13**, 227–243.

Hassanin A. & Ropiquet A. (2004) Molecular phylogeny of the tribe Bovini (Bovidae, Bovinae) and the taxonomic status of the Kouprey, *Bos sauveli* Urbain 1937. *Molecular Phylogenetics and Evolution*, **33**, 896–907.

Heads M. (2002) Birds of paradise, vicariance biogeography and terrane tectonics in New Guinea. *Journal of Biogeography*, **28**, 261–283.

Heads M. (2005) Dating nodes on molecular phylogenies: a critique of molecular biogeography. *Cladistics*, **21**, 62–78.

Heads M. (2009) Inferring biogeographic history from molecular phylogenies. *Biological Journal of the Linnean Society*, **98**, 757–774.

Heads M. (2011) Old taxa on young islands: a critique of the use of island age to date island-endemic clades and calibrate phylogenies. *Systematic Biology*, **60**, 204–218.

Heads M. (2012) Bayesian transmogrification of clade divergence dates: a critique. *Journal of Biogeography*, **39**, 1749–1756.

Heaney L.R. (1991) A synopsis of climatic and vegetational change in Southeast Asia. *Climatic Change*, **19**, 53–61.

Heaney L.R. (2007) Is a new paradigm emerging for oceanic island biogeography? *Journal of Biogeography*, **34**, 753–757.

Heinicke M.P., Duellman W.E. & Hedges S.B. (2007) Major Caribbean and Central American frog faunas originated by ancient oceanic dispersal. *Proceedings of the National Academy of Sciences USA*, **104**, 10092–10097.

Heled J. & Drummond A.J. (2010) Bayesian inference of species trees from multilocus data. *Molecular Biology and Evolution*, **27**, 570–580.

Herder F., Nolte A.W., Pfaender J., Schwarzer J., Hadiaty R.K. & Schliewen U.K. (2006a) Adaptive radiation and hybridization in Wallace's dreamponds: evidence from sailfin silversides in the Malili lakes of Sulawesi. *Proceedings of the Royal Society of London B*, **273**, 2209–2217.

Herder F. & Schliewen U.K. (2010) Beyond sympatric speciation: radiation of sailfin silverside fishes in the Malili lakes (Sulawesi). *Evolution in Action* (ed. by M. Glaubrecht), pp. 465–483. Springer, Berlin.

Herder F., Schwarzer J., Pfaender J., Hadiaty R.K. & Schliewen U.K. (2006b) Preliminary checklist of sailfin silversides (Teleostei: Telmatherinidae) in the Malili Lakes of Sulawesi (Indonesia), with a synopsis of systematics and threats. *Verhandlungen der Gesellschaft für Ichthyologie*, **5**, 139–163.

Hertwig S.T., Schweizer M., Das I. & Haas A. (2013) Diversification in a biodiversity hotspot – the evolution of Southeast Asian rhacophorid tree frogs on Borneo (Amphibia: Anura: Rhacophoridae). *Molecular Phylogenetics and Evolution*, **68**, 567–581.

Hill K.C. & Hall R. (2003) Mesozoic-Cenozoic evolution of Australia's New Guinea margin in a west Pacific context. *Geological Society of Australia, Special Publication/Geological Society of America, Special Papers*, **22/372**, 265–289.

Hipsley C.A. & Müller J. (2014) Beyond fossil calibrations: realities of molecular clock practices in evolutionary biology. *Frontiers in Genetics*, **5**, 138.

Ho S.Y.W. (2007) Calibrating molecular estimates of substitution rates and divergence times in birds. *Journal of Avian Biology*, **38**, 409–414.

Ho S.Y.W. (2014) The changing face of the molecular evolutionary clock. *Trends in Ecology & Evolution*, **29**, 496–503.

Ho S.Y.W. & Philipps M.J. (2009) Accounting for calibration uncertainty in phylogenetic estimation of evolutionary divergence times. *Systematic Biology*, **58**, 367–380.

Ho S.Y.W., Saarma U., Barnett R., Haile J. & Shapiro B. (2008) The effect of inappropriate calibration: three case studies in molecular ecology. *PLoS ONE*, **3**, e1615.

Ho S.Y.W. & Shapiro B. (2011) Skyline-plot methods for estimating demographic history from nucleotide sequences. *Molecular Ecology Resources*, **11**, 423–434.

Holloway J.D. (1990) Sulawesi biogeography – discussion and summing up. *Insects and the rain forests of South East Asia (Wallacea)* (ed. by W.J. Knight and J.D. Holloway), pp. 95–102. Royal Entomological Society, London.

Holt B.G., Lessard J.-P., Borregaard M.K., Fritz S.A., Araujo M.B., Dimitrov D., Fabre P.-H., Graham C.H., Graves G.R., Jonsson K.A., Nogues-Bravo D., Wang Z., Whittaker R.J., Fjeldsa J. & Rahbek C. (2013) An update of Wallace's zoogeographic regions of the world. *Science*, **339**, 74–78.

Hooijer D.A. (1952) *Palaeoloxodon* cf. *namadicus* (Falconer et Cautley) from Borneo. *Proceedings of the Koninklijke Nederlandse Akademie van Wetenschappen Series B*, **55**, 395–398.

Hoorn C., Wesselingh F.P., ter Steege H., Bermudez M.A., Mora A., Sevink J., Sanmartín I., Sanchez-Meseguer A., Anderson C.L., Figueiredo J.P., Jaramillo C., Riff D., Negri F.R., Hooghiemstra H., Lundberg J., Stadler T., Särkinen T. & Antonelli A. (2010) Amazonia through time: Andean uplift, climate change, landscape evolution, and biodiversity. *Science*, **330**, 927–931.

Hunn C.A. & Upchurch P. (2001) The importance of time/space in diagnosing the causality of phylogenetic events: towards a "chronobiogeographical" paradigm? *Systematic Biology*, **50**, 391–407.

Hunt T., Bergsten J., Levkanicova Z., Papadopoulou A., John O.S., Wild R., Hammond P.M., Ahrens D., Balke M., Caterino M.S., Gómez-Zurita J., Ribera I., Barraclough T.G., Bocakova M., Bocak L. & Vogler A.P. (2007) A comprehensive phylogeny of beetles reveals the evolutionary origins of a superradiation. *Science*, **318**, 1913–1916.

Huxley T.H. (1868) On the classification and distribution of the Alectoromorphae and Heteromorphae. *Proceedings of the Zoological Society of London*, 294–319.

Inger R.F. & Voris H.K. (2001) The biogeographical relations of the frogs and snakes of Sundaland. *Journal of Biogeography*, **28**, 863–891.

Jenkins P.D., Kilpatrick C.W., Robinson M.F. & Timmins R.J. (2005) Morphological and molecular investigations of a new family, genus and species of rodent (Mammalia: Rodentia: Hystricognatha) from Lao PDR. *Systematics and Biodiversity*, **2**, 419–454.

Jepson P. & Whittaker R.J. (2002) Ecoregions in context: a critique with special reference to Indonesia. *Conservation Biology*, **16**, 42–57.

Jønsson K.A., Fabre P.H., Ricklefs R.E. & Fjeldså J. (2011) Major global radiation of corvoid birds originated in the proto-Papuan archipelago. *Proceedings of the National Academy of Sciences USA*, **108**, 2328–2333.

Joyce D.A., Lunt D.H., Bills R., Turner G.F., Katongo C., Duftner N., Sturmbauer C. & Seehausen O. (2005) An extant cichlid fish radiation emerged in an extinct Pleistocene lake. *Nature*, **435**, 90–95.

Kass R.E. & Raftery A.E. (1995) Bayes factors. *Journal of the American Statistical Association*, **90**, 773–795.

Katoh K. & Toh H. (2008) Recent developments in the MAFFT multiple sequence alignment program. *Briefings in Bioinformatics*, **9**, 286–298.

Keast A. (2001) The vertebrate fauna of the Wallacean Island Interchange Zone: the basis of inbalance and impoverishment. *Faunal and floral migrations and evolution in SE Asia-Australasia* pp. 287–310. Balkema, Lisse.

Kikkawa Y., Yonekawa H., Suzuki H. & Amano T. (1997) Analysis of genetic diversity of domestic water buffaloes and anoas based on variations in the mitochondrial gene for cytochrome b. *Animal Genetics*, **28**, 195–201.

Klaus S., Schubart C.D., Streit B. & Pfenninger M. (2010) When Indian crabs were not yet Asian – biogeographic evidence for Eocene proximity of India and Southeast Asia. *BMC Evolutionary Biology*, **10**, 287.

Klaus S., Selvandran S., Goh J.W., Wowor D., Brandis D., Koller P., Schubart C.D., Streit B., Meier R., Ng P.K.L. & Yeo D.C.J. (2013) Out of Borneo: Neogene diversification of Sundaic freshwater crabs (Crustacea: Brachyura: Gecarcinucidae: *Parathelphusa*). *Journal of Biogeography*, **40**, 63–74.

Knapp M., Mudaliar R., Havell D., Wagstaff S.J. & Lockhart P.J. (2007) The drowning of New Zealand and the problem of *Agathis*. *Systematic Biology*, **56**, 862–870.

Knapp M., Stöckler K., Havell D., Delsuc F., Sebastiani F. & Lockhart P.J. (2005) Relaxed molecular clock provides evidence for long-distance dispersal of Nothofagus (southern beech). *PLoS Biology*, **3**, e14.

Kocher T.D., Conroy J.A., McKaye K.R., Stauffer J.R. & Lockwood S.F. (1995) Evolution of NADH dehydrogenase subunit 2 in East African cichlid fish. *Molecular Phylogenetics and Evolution*, **4**, 420–432.

Kodandaramaiah U. (2011) Tectonic calibrations in molecular dating. *Current Zoology*, **57**, 116–124.

Koh L.P. & Sodhi N.S. (2010) Conserving Southeast Asia's imperiled biodiversity: scientific, management, and policy challenges. *Biodiversity and Conservation*, **19**, 913–917.

Köhler F. & Dames C. (2009) Phylogeny and systematics of the Pachychilidae of mainland Southeast Asia – novel insights from morphology and mitochondrial DNA (Mollusca, Caenogastropoda, Cerithioidea). *Zoological Journal of the Linnean Society*, **157**, 679–699.

Köhler F. & Glaubrecht M. (2010) Uncovering an overlooked radiation: molecular phylogeny and biogeography of Madagascar's endemic river snails (Caenogastropoda: Pachychilidae: *Madagasikara* gen. nov.). *Biological Journal of the Linnean Society*, **99**, 867–894.

Köhler F., von Rintelen T., Meyer A. & Glaubrecht M. (2004) Multiple origin of viviparity in Southeast Asian gastropods (Cerithoidea: Pachychilidae) and its evolutionary implications. *Evolution*, **58**, 2215–2226.

Kottelat M. (1990) Sailfin silversides (Pisces: Telmatherinidae) of Lakes Towuti, Mahalona and Wawontoa (Sulawesi, Indonesia) with descriptions of two new genera and two new species. *Ichthyological Exploration of Freshwaters*, **1**, 35–54.

Kreft H. & Jetz W. (2007) Global patterns and determinants of vascular plant diversity. *Proceedings of the National Academy of Sciences USA*, **104**, 5925–5930.

Kreft H. & Jetz W. (2013) Comment on "An update of Wallace's zoogeographic regions of the world". *Science*, **341**, 343.

Larson G., Cucchi T., Fujita M., Matisoo-Smith E., Robins J., Anderson A., Rolett B., Spriggs M., Dolman G., Kim T.H., Thuy N.T.D., Randi E., Doherty M., Due R.A., Bollt R., Djubiantono T., Griffin B., Intoh M., Keane E., Kirch P., Li K.T., Morwood M., Pedriña L.M., Piper P.J., Rabett R.J., Shooter P., Van den Bergh G., West E., Wickler S., Yuan J., Cooper A. & Dobney K. (2007) Phylogeny and ancient DNA of *Sus* provides insights into neolithic expansion in Island Southeast Asia and Oceania. *Proceedings of the National Academy of Sciences USA*, **104**, 4834–4839.

Larson G., Dobney K., Albarella U., Fang M., Matisoo-Smith E., Robins J., Lowden S., Finlayson H., Brand T., Willerslev E., Rowley-Conwy P., Andersson L. & Cooper A. (2005) Worldwide phylogeography of wild boar reveals multiple centers of pig domestication. *Science*, **307**, 1618–1621.

Lerner H.R.L., Meyer M., James H.F., Hofreiter M. & Fleischer R.C. (2011) Multilocus resolution of phylogeny and timescale in the extant adaptive radiation of Hawaiian honeycreepers. *Current Biology*, **21**, 1838–1844.

Li J.-T., Li Y., Klaus S., Rao D.-Q., Hillis D.M. & Zhang Y.-P. (2013) Diversification of rhacophorid frogs provides evidence for accelerated faunal exchange between India and Eurasia during the Oligocene. *Proceedings of the National Academy of Sciences USA*, **110**, 3441–3446.

Li W.H., Tanimura M. & Sharp P.M. (1987) An evaluation of the molecular clock hypothesis using mammalian DNA sequences. *Journal of Molecular Evolution*, **25**, 330–342.

Librado P. & Rozas J. (2009) DnaSP v5: a software for comprehensive analysis of DNA polymorphism data. *Bioinformatics (Oxford, England)*, **25**, 1451–2.

Linkem C.W., Brown R.M., Siler C.D., Evans B.J., Austin C.C., Iskandar D.T., Diesmos A.C., Supriatna J., Andayani N. & McGuire J.A. (2013) Stochastic faunal exchanges drive diversification in widespread Wallacean and Pacific island lizards (Squamata: Scincidae: *Lamprolepis smaragdina*). *Journal of Biogeography*, **40**, 507–520.

Lisiecki L.E. & Raymo M.E. (2005) A Pliocene-Pleistocene stack of 57 globally distributed benthic $\delta^{18}O$ records. *Paleoceanography*, **20**, PA1003.

Liu L. & Pearl D.K. (2007) Species trees from gene trees: reconstructing Bayesian posterior distributions of a species phylogeny using estimated gene tree distributions. *Systematic Biology*, **56**, 504–514.

Lohman D.J., de Bruyn M., Page T., von Rintelen K., Hall R., Ng P.K.L., Shih H.T., Carvalho G.R. & von Rintelen T. (2011) Biogeography of the Indo-Australian archipelago. *Annual Review of Ecology, Evolution and Systematics*, **42**, 205–226.

Losos J.B. & Mahler D.L. (2010) Adaptive radiation: the interaction of ecological opportunity, adaptation, and speciation. *Evolution since Darwin: the first 150 years* (ed. by M.A. Bell, D.J. Futuyma, W.F. Eanes, and J.S. Levinton), pp. 381–420. Sinauer Associates, Sunderland, MA.

Louys J. & Meijaard E. (2010) Palaeoecology of Southeast Asian megafauna-bearing sites from the Pleistocene and a review of environmental changes in the region. *Journal of Biogeography*, **37**, 1432–1449.

Lukoschek V., Scott Keogh J. & Avise J.C. (2012) Evaluating fossil calibrations for dating phylogenies in light of rates of molecular evolution: a comparison of three approaches. *Systematic Biology*, **61**, 22–43.

Lydekker R. (1896) *A geographical history of mammals.* Cambridge University Press, Cambridge.

Maekawa K., Kon M., Araya K. & Matsumoto T. (2001) Phylogeny and biogeography of wood-feeding cockroaches, genus *Salganea* Stål (Blaberidae: Panesthiinae), in Southeast Asia based on mitochondrial DNA sequences. *Journal of Molecular Evolution*, **53**, 651–659.

Marivaux L., Bocat L., Chaimanee Y., Jaeger J.-J., Marandat B., Srisuk P., Tafforeau P., Yamee C. & Welcomme J.-L. (2006) Cynocephalid dermopterans from the Palaeogene of South Asia (Thailand, Myanmar and Pakistan): systematic, evolutionary and palaeobiogeographic implications. *Zoologica Scripta*, **35**, 395–420.

Marivaux L., Chaimanee Y., Yamee C., Srisuk P. & Jaeger J.-J. (2004) Discovery of *Fallomus ladakhensis* Nanda & Sahni, 1998 (Rodentia, Diatomyidae) in the lignites of Nong Ya Plong (Phetchaburi Province, Thailand): systematic, biochronological and paleoenvironmental implications. *Geodiversitas*, **26**, 493–507.

Martin C.H. & Wainwright P.C. (2013) Multiple fitness peaks on the adaptive landscape drive adaptive radiation in the wild. *Science*, **339**, 208–211.

Matzke N.J. (2013) Probabilistic historical biogeography: new models for founder-event speciation, imperfect detection, and fossils allow improved accuracy and model-testing. *Frontiers of Biogeography*, **5**, 242–248.

May R.M. (1990) How many species? *Philosophical Transactions of the Royal Society of London B*, **330**, 293–304.

Mayr E. (1944) Wallace's line in the light of recent zoogeographic studies. *Quarterly Review of Biology*, **19**, 1–14.

McGuigan K. (2000) Phylogenetic relationships and historical biogeography of melanotaeniid fishes in Australia and New Guinea. *Marine and Freshwater Research*, **51**, 713–723.

Meijer H.J.M. (2014) The avian fossil record in Insular Southeast Asia and its implications for avian biogeography and palaeoecology. *PeerJ*, **2**, e295.

Mercer J.M. & Roth V.L. (2003) The effects of cenozoic global change on squirrel phylogeny. *Science*, **299**, 1568–1572.

Merker S., Driller C., Perwitasari-Farajallah D., Pamungkas J. & Zischler H. (2009) Elucidating geological and biological processes underlying the diversification of Sulawesi tarsiers. *Proceedings of the National Academy of Sciences USA*, **106**, 8459–8464.

Metcalfe I. (2011a) Tectonic framework and Phanerozoic evolution of Sundaland. *Gondwana Research*, **19**, 3–21.

Metcalfe I. (2011b) Palaeozoic-Mesozoic history of SE Asia. *The Southeast Asian Gateway: history and tectonics of Australia-Asia collision* (ed. by R. Hall, M.A. Cottam, and M.E.J. Wilson), pp. 7–35. The Geological Society of London, London.

Michalak I., Zhang L.B. & Renner S.S. (2010) Trans-Atlantic, trans-Pacific and trans-Indian Ocean dispersal in the small Gondwanan Laurales family Hernandiaceae. *Journal of Biogeography*, **37**, 1214–1226.

Michaux B. (1991) Distributional patterns and tectonic development in Indonesia: Wallace reinterpreted. *Australian Systematic Botany*, **4**, 25–36.

Michaux B. (1994) Land movements and animal distributions in east Wallacea (eastern Indonesia, Papua New Guinea and Melanesia). *Palaeogeography, Palaeoclimatology, Palaeoecology*, **112**, 323–343.

Michaux B. (1996) The origin of southwest Sulawesi and other Indonesian terranes: a biological view. *Palaeogeography, Palaeoclimatology, Palaeoecology*, **122**, 167–183.

Michaux B. (2010) Biogeology of Wallacea: geotectonic models, areas of endemism, and natural biogeographic units. *Biological Journal of the Linnean Society*, **101**, 193–212.

Miettinen J., Shi C. & Liew S.C. (2011) Deforestation rates in insular Southeast Asia between 2000 and 2010. *Global Change Biology*, **17**, 2261–2270.

Miller K.G., Kominz M.A., Browning J. V, Wright J.D., Mountain G.S., Katz M.E., Sugarman P.J., Cramer B.S., Christie-Blick N. & Pekar S.F. (2005) The Phanerozoic record of global sea-level change. *Science*, **310**, 1293–1298.

Miller M.A., Pfeiffer W. & Schwartz T. (2010) Creating the CIPRES Science Gateway for inference of large phylogenetic trees. Proceedings of the Gateway Computing Environments Workshop (GCE), 1–8.

Milner-Gulland E.J., Bennett E.L. & SCB-2002-Annual-Meeting-Wild-Meat-Group (2003) Wild meat: the bigger picture. *Trends in Ecology & Evolution*, **18**, 351–357.

Mitchell K.J., Llamas B., Soubrier J., Rawlence N.J., Worthy T.H., Wood J., Lee M.S.Y. & Cooper A. (2014) Ancient DNA reveals elephant birds and kiwi are sister taxa and clarifies ratite bird evolution. *Science*, **344**, 898–900.

Moran M.D. (2003) Arguments for rejecting the sequential Bonferroni in ecological studies. *Oikos*, **100**, 403–405.

Morley R.J. (1996) Biostratigraphic characterization of systems tracts in Tertiary sedimentary basins. *Proceedings of the International Symposium on Sequence Stratigraphy in S.E. Asia*, 49–71.

Morley R.J. (2000) *Origin and evolution of tropical rain forests.* John Wiley & Sons, Chichester.

Morley R.J. (2006) Cretaceous and Tertiary climate change and the past distribution of megathermal rain forests. *Tropical rainforest responses to climatic change* (ed. by M. Bush and F. J.R.), pp. 1–54. Springer Praxis Books, Heidelberg.

Morley R.J. (2012) A review of the Cenozoic palaeoclimate history of Southeast Asia. *Biotic evolution and environmental change in Southeast Asia* (ed. by D. Gower, K.G. Johnson, B.R. Rosen, J. Richardson, L. Rüber, and S.T. Williams), pp. 79–114. Cambridge University Press, Cambridge.

Morley R.J. & Morley H.P. (2011) Neogene climate history of the Makassar Straits. *The Southeast Asian Gateway: history and tectonics of Australia-Asia collision* pp. 319–332. Geological Society of London, Special Publication, London.

Morley R.J. & Morley H.P. (2013) Mid Cenozoic freshwater wetlands of the Sunda region. *Journal of Limnology*, **72**, 18–35.

Morley R.J., Swiecicki T. & Pham D.T.T. (2012) A sequence stratigraphic framework for the Sunda region, based on integration of biostratigraphic, lithological and seismic data from Nam Con Son Basin, Vietnam. *Proceedings Indonesian Petroleum Association*, **35th Annual**, IPA11–G–002.

Morrone J.J. (2005) Cladistic biogeography: identity and place. *Journal of Biogeography*, **32**, 1281–1284.

Morrone J.J. & Crisci J.V. (1995) Historical biogeography: introduction to methods. *Annual Review of Ecology and Systematics*, **26**, 373–401.

Moss S.J. & Wilson M.E.J. (1998) Biogeographic implications of the Tertiary palaeogeographic evolution of Sulawesi and Borneo. *Biogeography and Geological Evolution of SE Asia* (ed. by R. Hall and J.D. Holloway), pp. 133–163. Backhuys Publishers, Leiden.

Moulds P.J. (1989) Development of the Bengkalis Depression, Central Sumatra and its subsequent deformation – a model for other Sumatran grabens? *18th Annual Convention Proceedings (Indones. Pet. Assoc.)*, 217–245.

Müller C.J. & Beheregaray L.B. (2010) Palaeo island-affinities revisited – Biogeography and systematics of the Indo-Pacific genus *Cethosia* Fabricius (Lepidoptera: Nymphalidae). *Molecular Phylogenetics and Evolution*, **57**, 314–326.

Müller C.J., Matos-Maraví P.F. & Beheregaray L.B. (2013) Delving into *Delias* Hübner (Lepidoptera: Pieridae): fine-scale biogeography, phylogenetics and systematics of the world's largest butterfly genus. *Journal of Biogeography*, **40**, 881–893.

Müller C.J., Wahlberg N. & Beheregaray L.B. (2010) "After Africa": the evolutionary history and systematics of the genus *Charaxes* Ochsenheimer (Lepidoptera: Nymphalidae) in the Indo-Pacific region. *Biological Journal of the Linnean Society*, **100**, 457–481.

Musser G.G., Smith A.L., Robinson M.F. & Lunde D.P. (2005) Description of a new genus and species of rodent (Murinae, Muridae, Rodentia) from the Khammouan Limestone National Biodiversity Conservation Area in Lao PDR. *American Museum Novitates*, **3497**, 1–32.

Myers N., Mittermeier R.A., Mittermeier C.G., Fonseca G.A.B. d & Kent J. (2000) Biodiversity hotspots for conservation priorities. *Nature*, **403**, 853–858.

Nattier R., Robillard T., Desutter-Grandcolas L., Couloux A. & Grandcolas P. (2011) Older than New Caledonia emergence? A molecular phylogenetic study of the eneopterine crickets (Orthoptera: Grylloidea). *Journal of Biogeography*, **38**, 2195–2209.

Near T.J. (2004) Estimating divergence times of notothenioid fishes using a fossil-calibrated molecular clock. *Antarctic Science*, **16**, 37–44.

Near T.J., Dornburg A., Eytan R.I., Keck B.P., Smith W.L., Kuhn K.L., Moore J.A., Price S.A., Burbrink F.T., Friedman M. & Wainwright P.C. (2013) Phylogeny and tempo of diversification in the superradiation of spiny-rayed fishes. *Proceedings of the National Academy of Sciences USA*, **110**, 12738–12743.

Nelson G. & Ladiges P.Y. (2009) Biogeography and the molecular dating game: a futile revival of phenetics? *Bulletin de la Societe géologique de France*, **180**, 39–43.

Obbard D.J., Maclennan J., Kim K.-W., Rambaut A., O'Grady P.M. & Jiggins F.M. (2012) Estimating divergence dates and substitution rates in the *Drosophila* phylogeny. *Molecular Biology and Evolution*, **29**, 3459–3473.

Ødegaard F. (2000) How many species of arthropods? Erwin's estimate revised. *Biological Journal of the Linnean Society*, **71**, 583–597.

Oliveros C.H. & Moyle R.G. (2010) Origin and diversification of Philippine bulbuls. *Molecular Phylogenetics and Evolution*, **54**, 822–832.

Parent C.E., Caccone A. & Petren K. (2008) Colonization and diversification of Galápagos terrestrial fauna: a phylogenetic and biogeographical synthesis. *Philosophical Transactions of the Royal Society of London B*, **363**, 3347–61.

Parenti L.R. (1993) Relationships of atherinomorph fishes (Teleostei). *Bulletin of Marine Science*, **52**, 170–196.

Parenti L.R. (1996) Phylogenetic systematics and biogeography of phallostethid fishes (Atherinomorpha, Phallostethidae) of Northwestern Borneo, with description of a new species. *Copeia*, **1996**, 703–712.

Parenti L.R. & Ebach M.C. (2010) Wallacea deconstructed. *Beyond cladistics: the branching of a paradigm* (ed. by D.M. Williams and S. Knapp), pp. 303–318.

Parenti L.R. & Ebach M.C. (2013) Evidence and hypothesis in biogeography. *Journal of Biogeography*, **40**, 813–820.

Parenti L.R. & Louie K.D. (1998) *Neostethus djajaorum*, new species, from Sulawesi, Indonesia, the first phallostethid fish (Teleostei: Atherinomorpha) known from east of Wallace's line. *Raffles Bulletin of Zoology*, **43**, 139–150.

Parham J.F., Donoghue P.C.J., Bell C.J., Calway T.D., Head J.J., Holroyd P.A., Inoue J.G., Irmis R.B., Joyce W.G., Ksepka D.T., Patané J.S.L., Smith N.D., Tarver J.E., van Tuinen M., Yang Z., Angielczyk K.D., Greenwood J.M., Hipsley C.A., Jacobs L., Makovicky P.J., Müller J., Smith K.T., Theodor J.M., Warnock R.C.M. & Benton M.J. (2012) Best practices for justifying fossil calibrations. *Systematic Biology*, **61**, 346–59.

Parr C.S., Guralnick R., Cellinese N. & Page R.D.M. (2012) Evolutionary informatics: unifying knowledge about the diversity of life. *Trends in Ecology & Evolution*, **27**, 94–103.

Peigné S., Chaimanee Y., Yamee C., Srisuk P., Marandat B. & Jaeger J.-J. (2006) A new member of the Mustelida (Mammalia: Carnivora) from the Paleogene of Southern Asia. *Journal of Vertebrate Paleontology*, **26**, 788–793.

Pereira S.L. & Baker A.J. (2006) A molecular timescale for galliform birds accounting for uncertainty in time estimates and heterogeneity of rates of DNA substitutions across lineages and sites. *Molecular Phylogenetics and Evolution*, **38**, 499–509.

Perneger T.V. (1998) What's wrong with Bonferroni adjustments. *BMJ*, **316**, 1236–1238.

Pesole G., Gissi C., de Chirico A. & Saccone C. (1999) Nucleotide substitution rate of mammalian mitochondrial genomes. *Journal of Molecular Evolution*, **48**, 427–434.

Pfaender J., Gray S.M., Rick I.P., Chapuis S., Hadiaty R.K. & Herder F. (2013) Spectral data reveal unexpected cryptic colour polymorphism in female sailfin silverside fish from ancient Lake Matano. *Hydrobiologia*, in press.

Pfaender J., Miesen F.W., Hadiaty R.K. & Herder F. (2011) Adaptive speciation and sexual dimorphism contribute to diversity in form and function in the adaptive radiation of Lake Matano's sympatric roundfin sailfin silversides. *Journal of Evolutionary Biology*, **24**, 2329–2345.

Pfaender J., Schliewen U.K. & Herder F. (2010) Phenotypic traits meet patterns of resource use in the radiation of "sharpfin" sailfin silverside fish in Lake Matano. *Evolutionary Ecology*, **24**, 957–974.

Phillips M.J., Page T.J., de Bruyn M., Huey J.A., Humphreys W.F., Hughes J.M., Santos S.R., Schmidt D.J. & Waters J.M. (2013) The linking of plate tectonics and evolutionary divergence. *Current Biology*, **23**, R603–R605.

Pons J., Ribera I., Bertranpetit J. & Balke M. (2010) Nucleotide substitution rates for the full set of mitochondrial protein-coding genes in Coleoptera. *Molecular Phylogenetics and Evolution*, **56**, 796–807.

Posada D. (2008) jModelTest: Phylogenetic model averaging. *Molecular Biology and Evolution*, **25**, 1253–1256.

Price J.P. & Clague D.A. (2002) How old is the Hawaiian biota? Geology and phylogeny suggest recent divergence. *Proceedings of the Royal Society of London B*, **269**, 2429–2435.

Pulquério M.J.F. & Nichols R.A. (2007) Dates from the molecular clock: how wrong can we be? *Trends in Ecology & Evolution*, **22**, 180–184.

Rambaut A. & Drummond A.J. (2007) Tracer v. 1.5. Available from http://tree.bio.ed.ac.uk/software/tracer

Redford K.H. (1992) The empty forest. *BioScience*, **42**, 412–422.

Ree R.H., Moore B.R., Webb C.O. & Donoghue M.J. (2005) A likelihood framework for inferring the evolution of geographic range on phylogenetic trees. *Evolution*, **59**, 2299–2311.

Ree R.H. & Smith S.A. (2008) Maximum likelihood inference of geographic range evolution by dispersal, local extinction, and cladogenesis. *Systematic Biology*, **57**, 4–14.

Ribas C.C., Aleixo A., Nogueira A.C.R., Miyaki C.Y. & Cracraft J. (2012) A palaeobiogeographic model for biotic diversification within Amazonia over the past three million years. *Proceedings of the Royal Society of London B*, **279**, 681–689.

Richardson J.E., Costion C.M. & Muellner A.N. (2012) The Malesian floristic interchange: plant migration patterns across Wallace's Line. *Biotic evolution and environmental change in Southeast Asia* (ed. by D. Gower, K.G. Johnson, B.R. Rosen, J. Richardson, L. Rüber, and S.T. Williams), pp. 138–163. Cambridge University Press, Cambridge.

Roderick G.K. & Gillespie R.G. (1998) Speciation and phylogeography of Hawaiian terrestrial arthropods. *Molecular Ecology*, **7**, 519–531.

Rohland N., Reich D., Mallick S., Meyer M., Green R.E., Georgiadis N., Roca A.L. & Hofreiter M. (2010) Genomic DNA sequences from mastodon and woolly mammoth reveal deep speciation of forest and savanna elephants. *PLoS Biology*, **8**, e1000564.

Rondinini C., Di Marco M., Chiozza F., Santulli G., Baisero D., Visconti P., Hoffmann M., Schipper J., Stuart S.N., Tognelli M.F., Amori G., Falcucci A., Maiorano L. & Boitani L. (2011) Global habitat suitability models of terrestrial mammals. *Philosophical Transactions of the Royal Society of London B*, **366**, 2633–2641.

Ronquist F. & Huelsenbeck J.P. (2003) MrBayes 3: Bayesian phylogenetic inference under mixed models. *Bioinformatics*, **19**, 1572–1574.

Ronquist F. & Sanmartín I. (2011) Phylogenetic methods in biogeography. *Annual Review of Ecology, Evolution and Systematics*, **42**, 441–464.

Roy D., Paterson G., Hamilton P.B., Heath D.D. & Haffner G.D. (2007) Resource-based adaptive divergence in the freshwater fish *Telmatherina* from Lake Matano, Indonesia. *Molecular Ecology*, **16**, 35–48.

Rüber L., van Tassell J.L. & Zardoya R. (2003) Rapid speciation and ecological divergence in the American seven-spined gobies (Gobiidae, Gobiosomatini) inferred from a molecular phylogeny. *Evolution*, **57**, 1584–1598.

Ruedi M., Auberson M. & Savolainen V. (1998) Biogeography of Sulawesian shrews: testing for their origin with a parametric bootstrap on molecular data. *Molecular Phylogenetics and Evolution*, **9**, 567–571.

Ruedas L.A. & Morales J.C. (2005) Evolutionary relationships among genera of Phalangeridae (Metatheria: Diprotodontia) inferred from mitochondrial DNA. *Journal of Mammalogy*, **86**, 353–365.

Rutschmann F., Eriksson T., Salim K.A. & Conti E. (2007) Assessing calibration uncertainty in molecular dating: the assignment of fossils to alternative calibration points. *Systematic Biology*, **56**, 591–608.

Saeed B. & Ivantsoff W. (1991) *Kalyptatherina*, the first telmatherinid genus known outside of Sulawesi. *Ichthyological Exploration of Freshwaters*, **2**, 227–238.

Saeed B., Ivantsoff W. & Crowley L.E.L.M. (1994) Systematic relationships of atheriniform fishes within division 1 of the series Atherinomorpha (Actinopterygii) with relevant historical perspectives. *Voprosy Ichtiologii*, **34**, 1–32.

Samonds K.E., Godfrey L.R., Ali J.R., Goodman S.M., Vences M., Sutherland M.R., Irwin M.T. & Krause D.W. (2012) Spatial and temporal arrival patterns of Madagascar's vertebrate fauna explained by distance, ocean currents, and ancestor type. *Proceedings of the National Academy of Sciences USA*, **109**, 5352–5357.

Samonds K.E., Godfrey L.R., Ali J.R., Goodman S.M., Vences M., Sutherland M.R., Irwin M.T. & Krause D.W. (2013) Imperfect isolation: factors and filters shaping Madagascar's extant vertebrate fauna. *PLoS ONE*, **8**, e62086.

Sanders K.L. & Lee M.S.Y. (2007) Evaluating molecular clock calibrations using Bayesian analyses with soft and hard bounds. *Biology Letters*, **3**, 275–279.

Santos M.E. & Salzburger W. (2012) How cichlids diversify. *Science*, **338**, 619–621.

Sarasin P. & Sarasin F. (1901) *Ueber die geologische Geschichte der Insel Celebes auf Grund der Thierverbreitung.* Kreidel, Wiesbaden.

Sathiamurthy E. & Voris H.K. (2006) Maps of Holocene sea level transgression and submerged lakes on the Sunda Shelf. *The Natural History Journal of Chulalongkorn University*, **Suppl. 2**, 1–44.

Schluter D. (2000) *The ecology of adaptive radiation.* Oxford University Press, Oxford.

Schreiber A., Seibold I., Nötzold G. & Wink M. (1999) Cytochrome b gene hyplotypes characterize chromosomal lineages of Anoa, the Sulawesi dwarf buffalo (Bovidae: *Bubalus* sp.). *Journal of Heredity*, **90**, 165–176.

Schubart C.D., Diesel R. & Hedges S.B. (1998) Rapid evolution to terrestrial life in Jamaican crabs. *Nature*, **393**, 363–365.

Schubart C.D. & Ng P.K.L. (2008) A new molluscivore crab from Lake Poso confirms multiple colonization of ancient lakes in Sulawesi by freshwater crabs (Decapoda: Brachyura). *Zoological Journal of the Linnean Society*, **154**, 211–221.

Schwarzer J., Herder F., Misof B., Hadiaty R.K. & Schliewen U.K. (2008) Gene flow at the margin of Lake Matano's adaptive sailfin silverside radiation: Telmatherinidae of River Petea in Sulawesi. *Hydrobiologia*, **615**, 201–213.

Schwarzer J., Misof B., Ifuta S.N. & Schliewen U.K. (2011) Time and origin of cichlid colonization of the lower Congo rapids. *PLoS ONE*, **6**, e22380.

Schweizer M., Seehausen O., Güntert M. & Hertwig S.T. (2010) The evolutionary diversification of parrots supports a taxon pulse model with multiple trans-oceanic dispersal events and local radiations. *Molecular Phylogenetics and Evolution*, **54**, 984–994.

Setiadi M.I., McGuire J.A., Brown R.M., Zubairi M., Iskandar D.T., Andayani N., Supriatna J. & Evans B.J. (2011) Adaptive radiation and ecological opportunity in Sulawesi and Philippine fanged frog (*Limnonectes*) communities. *American Naturalist*, **178**, 221–240.

Setiamarga D.H.E., Miya M., Yamanoue Y., Azuma Y., Inoue J.G., Ishiguro N.B., Mabuchi K. & Nishida M. (2009) Divergence time of the two regional medaka populations in Japan as a new time scale for comparative genomics of vertebrates. *Biology Letters*, **5**, 812–816.

Setiamarga D.H.E., Miya M., Yamanoue Y., Mabuchi K., Satoh T.P., Inoue J.G. & Nishida M. (2008) Interrelationships of Atherinomorpha (medakas, flyingfishes, killifishes, silversides, and their relatives): the first evidence based on whole mitogenome sequences. *Molecular Phylogenetics and Evolution*, **49**, 598–605.

Shapiro B., Sibthorpe D., Rambaut A., Austin J., Wragg G.M., Bininda-Edmonds O.R.P., Lee P.L.M. & Cooper A. (2002) Flight of the Dodo. *Science*, **295**, 1683.

Shih H.T., Yeo D.C.J. & Ng P.K.L. (2009) The collision of the Indian plate with Asia: molecular evidence for its impact on the phylogeny of freshwater crabs (Brachyura: Potamidae). *Journal of Biogeography*, **36**, 703–719.

Shoup R.C., Morley R.J., Swiecicki T. & Clark S.J. (2012) Tectono-stratigraphic framework and Tertiary paleogeography of Southeast Asia: Gulf of Thailand to South Vietnam Shelf. *AAPG Search and Discovery article #30246 (posted Sept 24 2012)*, .

Siler C.D., Oaks J.R., Welton L.J., Linkem C.W., Swab J.C., Diesmos A.C. & Brown R.M. (2012) Did geckos ride the Palawan raft to the Philippines? *Journal of Biogeography*, **39**, 1217–1234.

Simpson G.G. (1977) Too many lines; the limits of the Oriental and Australian zoogeographic regions. *Proceedings of the American Philosophical Society*, **121**, 107–120.

Simpson G.G. (1980) *Splendid isolation: the curious history of South American mammals.* Yale University Press, New Haven.

Slik J.W.F., Aiba S.-I., Bastian M., Brearley F.Q., Cannon C.H., Eichhorn K.A.O., Fredriksson G., Kartawinata K., Laumonier Y., Mansor A., Marjokorpi A., Meijaard E., Morley R.J., Nagamasu H., Nilus R., Nurtjahya E., Payne J., Permana A., Poulsen A.D., Raes N., Riswan S., Van Schaik C.P., Sheil D., Sidiyasa K., Suzuki E., van Valkenburg J.L.C.H., Webb C.O., Wich S., Yoneda T., Zakaria R. & Zweifel N. (2011) Soils on exposed Sunda Shelf shaped biogeographic patterns in the equatorial forests of Southeast Asia. *Proceedings of the National Academy of Sciences USA*, **108**, 12343–12347.

Spakman W. & Hall R. (2010) Surface deformation and slab–mantle interaction during Banda arc subduction rollback. *Nature Geoscience*, **3**, 562–566.

Sparks J.S. & Smith W.L. (2004) Phylogeny and biogeography of the Malagasy and Australasian rainbowfishes (Teleostei: Melanotaenioidei): Gondwanan vicariance and evolution in freshwater. *Molecular Phylogenetics and Evolution*, **33**, 719–734.

Stamatakis A., Hoover P. & Rougemont J. (2008) A rapid bootstrap algorithm for the RAxML web servers. *Systematic Biology*, **57**, 758–771.

Stelbrink B., Albrecht C., Hall R. & von Rintelen T. (2012) The biogeography of Sulawesi revisited: is there evidence for a vicariant origin of taxa on Wallace's "anomalous island"? *Evolution*, **66**, 2252–2271.

Stresemann E. (1939) Die Vögel von Celebes. *Journal für Ornithologie*, **87**, 299–425.

Suchard M.A., Weiss R.E. & Sinsheimer J.S. (2001) Bayesian selection of continuous-time Markov chain evolutionary models. *Molecular Biology and Evolution*, **18**, 1001–1013.

Sun X. & Wang P. (2005) How old is the Asian monsoon system? – Palaeobotanical records from China. *Palaeogeography, Palaeoclimatology, Palaeoecology*, **222**, 181–222.

Swenson U., Nylinder S. & Munzinger J. (2014) Sapotaceae biogeography supports New Caledonia being an old Darwinian island. *Journal of Biogeography*, **41**, 797–809.

Swofford D.L. (2002) PAUP* (version 4.0). Phylogenetic analysis using parsimony (*and other methods).

Takehana Y., Naruse K. & Sakaizumi M. (2005) Molecular phylogeny of the medeka fishes genus *Oryzias* (Beloniformes: Adrianichthyidae) based on nuclear and mitochondrial DNA sequences. *Molecular Phylogenetics and Evolution*, **36**, 417–428.

Tamura K. & Nei M. (1993) Estimation of the number of nucleotide substitutions in the control region of mitochondrial DNA in humans and chimpanzees. *Molecular Biology and Evolution*, **10**, 512–526.

Thomas J.A., Welch J.J., Woolfit M. & Bromham L. (2006) There is no universal molecular clock for invertebrates, but rate variation does not scale with body size. *Proceedings of the National Academy of Sciences USA*, **103**, 7366–7371.

Thomson V.A., Lebrasseur O., Austin J.J., Hunt T.L., Burney D.A., Denham T., Rawlence N.J., Wood J.R., Gongora J., Girdland Flink L., Linderholm A., Dobney K., Larson G. & Cooper A. (2014) Using ancient DNA to study the origins and dispersal of ancestral Polynesian chickens across the Pacific. *Proceedings of the National Academy of Sciences USA*, **111**, 4826–4831.

Thornton I.W.B. (1983) Vicariance and dispersal: confrontation or compatibility? *GeoJournal*, **7**, 557–564.

Tosi A.J., Morales J.C. & Melnick D.J. (2003) Paternal, maternal, and biparental molecular markers provide unique windows onto the evolutionary history of macaque monkeys. *Evolution*, **57**, 1419–1435.

Turchetto-Zolet A.C., Pinheiro F., Salgueiro F. & Palma-Silva C. (2013) Phylogeographical patterns shed light on evolutionary process in South America. *Molecular Ecology*, **22**, 1193–1213.

Turner H., Hovenkamp P. & van Welzen P.C. (2001) Biogeography of Southeast Asia and the West Pacific. *Journal of Biogeography*, **28**, 217–230.

Unmack P.J. (2001) Biogeography of Australian freshwater fishes. *Journal of Biogeography*, **28**, 1053–1089.

Unmack P.J., Allen G.R. & Johnson J.B. (2013) Phylogeny and biogeography of rainbowfishes (Melanotaeniidae) from Australia and New Guinea. *Molecular Phylogenetics and Evolution*, **67**, 15–27.

Upchurch P. (2008) Gondwanan break-up: legacies of a lost world? *Trends in Ecology & Evolution*, **23**, 229–236.

Vaidya G., Lohman D.J. & Meier R. (2011) SequenceMatrix: concatenation software for the fast assembly of multi-gene datasets with character set and codon information. *Cladistics*, **27**, 171–180.

Vaillant J.J., Haffner G.D. & Cristecu M.E. (2011) The ancient lakes of Indonesia: towards integrated research on speciation. *Integrative and Comparative Biology*, **51**, 634–643.

Vakhrameev V.A. (1991) *Jurassic and Cretaceous floras and climates of the Earth.* Cambridge University Press, Cambridge.

van den Bergh G.D., de Vos J. & Sondaar P.Y. (2001) The late Quaternary palaeogeography of mammal evolution in the Indonesian archipelago. *Palaeogeography, Palaeoclimatology, Palaeoecology*, **171**, 385–408.

van den Bergh G.D., Meijer H.J.M., Awe R.D., Morwood M.J., Szabó K., van den Hoek Ostende L.W., Saptomo E.W., Piper P.J. & Dobney K.M. (2009) The Liang Bua faunal remains: a 95 k.yr. sequence from Flores, East Indonesia. *Journal of Human Evolution*, **57**, 527–537.

van Oosterzee P. (1997) *Where worlds collide. The Wallace Line.* Cornell University Press, Ithaca.

van Welzen P.C., Parnell J.A.N. & Slik J.W.F. (2011) Wallace's Line and plant distributions: two or three phytogeographical areas and where to group Java? *Biological Journal of the Linnean Society*, **103**, 531–545.

van Welzen P.C., Turner H. & Hovenkamp P. (2003) Historical biogeography of Southeast Asia and the West Pacific, or the generality of unrooted area networks as historical biogeographic hypotheses. *Journal of Biogeography*, **30**, 181–192.

Vences M., Vieites D.R., Glaw F., Brinkmann H., Kosuch J., Veith M. & Meyer A. (2003) Multiple overseas dispersal in amphibians. *Proceedings of the Royal Society of London B*, **270**, 2435–2442.

von Rintelen T., Bouchet P. & Glaubrecht M. (2007) Ancient lakes as hotspots of diversity: a morphological review of an endemic species flock of *Tylomelania* (Caenogastropoda: Cerithioidea: Pachychilidae) in the Malili lake system on Sulawesi, Indonesia. *Hydrobiologia*, **592**, 11–94.

von Rintelen T. & Glaubrecht M. (2005) Anatomy of an adaptive radiation: a unique reproductive strategy in the endemic freshwater gastropod *Tylomelania* (Cerithioidea: Pachychilidae) on Sulawesi, Indonesia and its biogeographical implications. *Biological Journal of the Linnean Society*, **85**, 513–542.

von Rintelen T. & Glaubrecht M. (2006) Rapid evolution of sessility in an endemic species flock of the freshwater bivalve *Corbicula* from ancient lakes on Sulawesi, Indonesia. *Biology Letters*, **2**, 73–77.

von Rintelen T. & Glaubrecht M. (2008) Three new species of the freshwater snail genus *Tylomelania* (Caenogastropoda: Pachychilidae) from the Malili lake system, Sulawesi, Indonesia. *Zootaxa*, **1852**, 37–49.

von Rintelen T., von Rintelen K. & Glaubrecht M. (2010) The species flocks of the viviparous freshwater gastropod Tylomelania (Mollusca: Cerithioidea: Pachychilidae) in the ancient lakes of Sulawesi, Indonesia: the role of geography, trophic morphology and color as driving forces in adaptive radiation. *Evolution in Action* (ed. by M. Glaubrecht), pp. 485–512. Springer, Berlin.

von Rintelen T., von Rintelen K., Glaubrecht M., Schubart C.D. & Herder F. (2012) Aquatic biodiversity hotspots in Wallacea: the species flocks in the ancient lakes of Sulawesi, Indonesia. *Biotic evolution and environmental change in Southeast Asia* (ed. by D.J. Gower, K.G. Johnson, J.E. Richardson, B.R. Rosen, L. Rüber, and S.T. Williams), pp. 290–315. Cambridge University Press, Cambridge.

von Rintelen T., Wilson A.B., Meyer A. & Glaubrecht M. (2004) Escalation and trophic specialization drive adaptive radiation of viviparous freshwater gastropods in the ancient lakes on Sulawesi, Indonesia. *Proceedings of the Royal Society of London B*, **271** , 2541–2549.

Voris H.K. (2000) Maps of Pleistocene sea levels in Southeast Asia: shorelines, river systems and time durations. *Journal of Biogeography*, **27**, 1153–1167.

Wakita K. & Metcalfe I. (2005) Ocean plate stratigraphy in East and Southeast Asia. *Journal of Asian Earth Sciences*, **24**, 679–702.

Wallace A.R. (1859) Letter from Mr. Wallace concerning the geographical distribution of birds. *Ibis*, **1**, 449–454.

Wallace A.R. (1860) On the zoological geography of the Malay Archipelago. *Journal of the Proceedings of the Linnean Society*, **4**, 172–184.

Wallace A.R. (1863) On the physical geography of the Malay Archipelago. *Journal of the Royal Geographical Society*, **33**, 217–234.

Wallace A.R. (1869) *The Malay Archipelago, the land of the orang-utan and the bird of paradise: a narrative of travel with studies of man and nature.* Macmillan, London.

Wallace A.R. (1876) *The geographical distribution of animals.* Harper & Brothers, New York.

Wallace A.R. (1880) *Island life, or the phenomena and causes of insular faunas and floras, including a revision and attempted solution of the problem of geological climates.* Macmillan, London.

Wallace A.R. (1910) *The world of life.* Chapman & Hall, London.

Wallis G.P. & Trewick S.A. (2009) New Zealand phylogeography: evolution on a small continent. *Molecular Ecology*, **18**, 3548–3580.

Walter R.P., Hogan J.D., Haffner G.D. & Heath D.D. (2011) Genetic structure and connectivity among lake populations of threatened *Paratherina* sailfin silversides from Sulawesi, Indonesia. *Conservation Genetics*, **12**, 1387–1393.

Walton C., Butlin R.K. & Monk K.A. (1997) A phylogeny for grasshoppers of the genus *Chitaura* (Orthoptera: Acrididae) from Sulawesi, Indonesia, based on mitochondrial DNA sequence data. *Biological Journal of the Linnean Society*, **62**, 365–382.

Warnock R.C.W., Yang Z. & Donoghue P.C.J. (2012) Exploring uncertainty in the calibration of the molecular clock. *Biology Letters*, **8**, 156–159.

Warren B.H., Strasberg D., Bruggemann J.H., Prys-Jones R.P. & Thébaud C. (2010) Why does the biota of the Madagascar region have such a strong Asiatic flavour? *Cladistics*, **26**, 526–538.

Waters J.M. & Craw D. (2006) Goodbye Gondwana? New Zeland biogeography, geology, and the problem of circularity. *Systematic Biology*, **55**, 351–356.

Waters J.M., Trewick S.A., Paterson A.M., Spencer H.G., Kennedy M., Craw D., Burridge C.P. & Wallis G.P. (2013) Biogeography off the tracks. *Systematic Biology*, **62**, 494–498.

Weir J.T. & Schluter D. (2008) Calibrating the avian molecular clock. *Molecular Ecology*, **17**, 2321–2328.

Wertheim J.O. & Sanderson M.J. (2010) Estimating diversification rates: how useful are divergence times? *Evolution*, **65**, 309–320.

Whitmore T.C. (1981) *Wallace's line and plate tectonics.* Clarendon Press, Oxford.

Whitmore T.C. (1987) *Biogeographical evolution of the Malay Archipelago.* Clarendon Press, Oxford.

Whitmore T.C. (1987) *Biogeographical Evolution of the Malay Archipelago.* Clarendon Press, Oxford.

Whittaker R.J. & Fernández-Palacios J.M. (2007) *Island biogeography. Ecology, evolution, and conservation.* Oxford University Press, Oxford.

Whitten A.J., Mustafa M. & Henderson G.S. (2002) *The ecology of Sulawesi.* Periplus, Singapore.

Wiens J.J. & Donoghue M.J. (2004) Historical biogeography, ecology and species richness. *Trends in Ecology & Evolution*, **19**, 639–644.

Wikramanayake E., Dinerstein E., Loucks C., Olson D., Morrison J., Lamoreux J., McKnight M. & Hedao P. (2002) Ecoregions in ascendance: reply to Jepson and Whittaker. *Conservation Biology*, **16**, 238–243.

Wilcove D.S., Giam X., Edwards D.P., Fisher B. & Koh L.P. (2013) Navjot's nightmare revisited: logging, agriculture, and biodiversity in Southeast Asia. *Trends in Ecology & Evolution*, **28**, 531–540.

Wiley E.O. (1988) Vicariance biogeography. *Annual Review of Ecology and Systematics*, **19**, 513–542.

Wilke T., Schultheiß R. & Albrecht C. (2009) As time goes by: a simple fool's guide to molecular clock approaches in invertebrates. *American Malacological Bulletin*, **27**, 25–45.

Willis S.C., Farias I.P. & Ortí G. (2014) Testing mitochondrial capture and deep coalescence in Amazonian cichlid fishes (Cichlidae: *Cichla*). *Evolution*, **68**, 256–268.

Wilson M.E.J. & Moss S.J. (1999) Cenozoic palaeogeographic evolution of Sulawesi and Borneo. *Palaeogeography, Palaeoclimatology, Palaeoecology*, **145**, 303–337.

Witts D., Hall R., Nichols G. & Morley R. (2012) A new depositional and provenance model for the Tanjung Formation, Barito Basin, SE Kalimantan, Indonesia. *Journal of Asian Earth Sciences*, **56**, 77–104.

Wong K.M. (2011) A biogeographic history of Southeast Asian rainforests. *Managing the future of Southeast Asia's valuable tropical rainforests* (ed. by R. Wickneswari and C. Cannon), pp. 21–55. Springer Netherlands, Dordrecht.

Woodruff D.S. & Turner L.M. (2009) The Indochinese-Sundaic zoogeographic transition: a description and analysis of terrestrial mammal species distributions. *Journal of Biogeography*, **36**, 803–821.

Woodruff D.S. (2010) Biogeography and conservation in Southeast Asia: how 2.7 million years of repeated environmental fluctuations affect today's patterns and the future of the remaining refugial-phase biodiversity. *Biodiversity and Conservation*, **19**, 919–941.

Xia X. & Xie Z. (2001) DAMBE: Data analysis in molecular biology and evolution. *Journal of Heredity*, **92**, 371–373.

Xia X., Xie Z., Salemi M., Chen L. & Wang Y. (2003) An index of substitution saturation and its application. *Molecular Phylogenetics and Evolution*, **26**, 1–7.

Yang Z. & Rannala B. (2006) Bayesian estimation of species divergence times under a molecular clock using multiple fossil calibrations with soft bounds. *Molecular Biology and Evolution*, **23**, 212–226.

Yoder A.D. & Nowak M.D. (2006) Has vicariance or dispersal been the predominant biogeographic force in Madagascar? Only time will tell. *Annual Review of Ecology, Evolution and Systematics*, **37**, 405–431.

Zachos J.C., Dickens G.R. & Zeebe R.E. (2008) An early Cenozoic perspective on greenhouse warming and carbon-cycle dynamics. *Nature*, **451**, 279–283.

Zachos J., Pagani M., Sloan L., Thomas E. & Billups K. (2001) Trends, rhythms, and aberrations in global climate 65 Ma to present. *Science*, **292**, 393–686.

Zuckerkandl E. (1987) On the molecular evolutionary clock. *Journal of Molecular Evolution*, **26**, 34–46.

Appendix

Supplementary Information – Chapter 1

Detailed Southeast Asian geological background

The various continental pieces that now make up SE Asia (Fig. S1) were all derived from the southern hemisphere supercontinent Gondwana (Metcalfe, 1988), and travelled north to progressively collide and coalesce prior to the current ongoing collision with the northwards moving Australian continent (Metcalfe & Irving, 1990; Metcalfe, 1996, 2011a, 2011b). During the separation of the various continental terranes from Gondwana, their northwards migration and collision, three intervening Tethyan oceans, the Paleo-Tethys, Meso-Tethys and Ceno-Tethys, were opened and subsequently destroyed. Remnants of these ancient oceans are preserved in the various narrow suture zones bounding the continental blocks, including ophiolitic rocks, volcanic arcs, and accretionary complexes with melange and deep sea sediments (Wakita & Metcalfe, 2005). The continental collisions in the SE Asia region occurred in two distinct phases. The earlier phase brought together pieces of continent that exhibit widely differing Late Paleozoic biotas representative of the high-latitude southern hemisphere Gondwana and low-latitude equatorial-northern hemisphere Cathaysian biotic provinces (Metcalfe, 2011b). The second collisional phase involved the addition of Gondwana and Asia-derived continental fragments to the mainland Sundaland core during the Mesozoic and Cenozoic and the Cenozoic collision of the northwards moving Australian continent with maritime SE Asia (Hall, 1996, 2002, 2011). The Late Paleozoic Gondwana-Cathaysia biogeographic divide or line in mainland Southeast Asia is as striking and as significant biogeographically as the extant Wallace/Huxley/Lydekker Lines that divide Australian and Asian biotas (Metcalfe, 2011a).

Long-lived plate tectonic convergence in the SE Asia region with long-term subduction has produced multiple volcanic arcs and island arc chains. The immense volume of volcanism must have impacted significantly on the atmosphere and at times driven climate oscillations both locally and globally (Hay, 1996). The opening of marginal ocean basins due to subduction processes, including subduction rollback, has also produced complex changes in SE Asian geomorphology and climate. Continent-continent collisions have resulted in mountain building events, major plutonism (e.g., tin-bearing granite belts of SE Asia), uplift and basin development. Newly formed mountains provide barriers to atmospheric circulation, trap water vapor, effect heat exchanges and affect seasonal climate extremes and monsoonal circulation. Short-lived islands (e.g., volcanoes atop volcanic arcs or over hotspots) will have aided dispersal events. SE Asia's biodiversity hotspots are rooted in the geological collisional history of the region.

Mesozoic and earlier

The western pre-Cenozoic core of SE Asia (Sundaland) comprises two continental blocks, a western Sibumasu Block and an eastern Indochina-East Malaya Block with a volcanic arc, the Sukhothai Arc sandwiched between these (Fig. S1) (Sone & Metcalfe, 2008; Metcalfe, 2011a, 2011b). The Sibumasu and Indochina-East Malaya Blocks form part of a collage of continental blocks (terranes), also including North and South China, Tarim and Qiangtang that were derived from Gondwana in the Paleozoic with the opening of the Paleo- and Meso-Tethys oceans (Fig. S2). The Sukhothai Arc was constructed on the margin of Indochina-East Malaya and then separated by back-arc spreading. These terranes were amalgamated by closure of the Paleo-Tethys and back-arc basin in the Triassic. A sliver-like Cathaysian continental terrane comprising West Sumatra and West Burma was added by strike-slip juxtaposition out board of Sibumasu, also in the Triassic (Barber & Crow, 2003, 2009; Metcalfe, 2005). Thus, by the end of the Triassic, core Sundaland, comprising Sumatra, the Thai-Malay peninsula and most of the present-day Sunda Shelf was part of continental Asia (Fig. S2).

During the Jurassic and Cretaceous, an ocean island arc developed along the southern margin of the Meso-Tethys by southwards subduction. This island arc has been referred to as the 'Woyla' Arc in the east and has been referred to as the 'Incertus' arc in the west (Hall, 2011). The 'Incertus' arc of Hall (2011) is here interpreted to be the Kohistan-Ladakh Arc. The Kohistan Arc may have formed as early as c. 135 Mya (Bosch et al., 2011) or even 150 Mya (Bouilhol et al., 2010b) and was equatorial in the Late Cretaceous-early Paleocene (Fig. S3) based on paleomagnetism (Khan et al., 2009). Arc magmatism ended by 61 Mya by collision with India and the arc was then carried forward with India to collide with Asia c. 50 Mya (Khan et al., 2009). Recent U-Pb/Hf/Nd isotopic data indicates an abrupt shift from juvenile isotopic signatures in the Jurassic-early Paleocene to evolved crustal like signatures in the Eocene (Bouilhol et al., 2010a, 2010b) supporting this contention.

Additional small continental blocks separated from the western Australian margin (Banda Embayment-Argo Abyssal Plain region) by back-arc rifting in the easternmost segment of the Kohistan-Woyla subduction zone system (Fig. S3). The block that separated from the Banda Embayment is here interpreted to be South West Borneo (also referred to as the Banda Block) following Hall et al. (2009) and Metcalfe (2011b). The block that separated from the Argo Abyssal Plain, previously referred to as Argoland, is now interpreted as East Java, West Sulawesi and possibly other small continental areas such as Bawean, Paternoster and Mangkalihat (Smyth et al., 2007; van Leeuwen et al., 2007; Metcalfe, 2011b). Southwest Borneo and Argoland migrated northwards as the Meso-Tethys narrowed and as Ceno-Tethys opened to the south (Fig. S3). Accretion of these blocks to Southeast Sundaland was completed by Late Cretaceous times.

Detailed Discussion on Development of the SE Asian Flora

This brief overview covers the region from the time of appearance of angiosperms, in the Early Cretaceous.

Angiosperm origins and the early Cretaceous flora
The suggestion that the 'area between Assam and Fiji' was the 'cradle of the angiosperms' (Takhtajan, 1969), driving perceptions of the early evolution of SE Asian flora for decades, has not been borne out by the fossil record, since palynofloras from the Early Cretaceous are of low diversity without the common angiosperm elements seen from the same time period in Western Gondwana (Morley, 2000). The diverse 'early angiosperm' elements seen in the modern flora, which prompted Takhtajan's hypothesis (Takhtajan, 1969) were subsequently shown to be relict, and due to a combination of global climate change and shifting tectonic plate positions (Morley, 2001). The claims for a Jurassic origin of angiosperms in China (Sun et al., 1998) were also shown to be invalid by Swisher III et al. (2002) due to erroneous dating.

Cretaceous to Paleocene vegetation
During the Early Cretaceous, the vegetation across the region was dominated by the gymnosperm family Cheirolepidiaceae, which produced the dispersed pollen *Classopollis* and *Exesipollenites,* suggesting a paleoclimate with some degree of water stress. Angiosperm elements subsequently increased in numbers during the later Cretaceous (Muller, 1968; Morley, 2000) but never attained the high diversities suggested by time-equivalent pollen floras in Central Africa (e.g., Boltenhagen, 1976) which are considered by Morley (2000, 2006) and Kubitzki (2005) to reflect the development of the world's first closed canopy rain forests. Paleocene palynomorph diversities remained low in the region of Borneo with most types having no modern analogue. However, the presence of *Nypa* pollen testifies to the presence of mangroves, and common ephedroids and Laurasian conifer pollen, suggest upland open woodland vegetation also subject to water stress. From the southern margin of Sundaland, from offshore East Java, common palm pollen from Paleocene or latest Cretaceous coaly sediments suggest a wetter climate in that area (Morley, 2012).

Eocene and the invasion of Indian floristic elements
By the earliest middle Eocene, the flora of the region had changed dramatically compared to that of the Paleocene, following the collision of the Indian Plate with Asia (Fig. S4, 45 Mya) and the dispersal of many taxa of Indian origin into the area, resulting in the extinction of many elements of the older Paleocene flora (Morley, 1998, 2000). The Indian immigrants included many elements, which today we consider particularly 'Southeast Asian' in character including the fabled *Durio* (Bombacaceae), *Gonystylus* (Gonystylaceae), many palms, including *Eugeissona* and Iguanurinae (Arecaceae) and ultimately Dipterocarpaceae. This flora was termed the 'Proto Indian' flora by Morley (2000). The Proto-Indian flora developed in stages between about 49 and 45 Mya, and by the middle Eocene, had diversified dramatically, reflected in a pollen flora with over 300 pollen types recorded from the middle Eocene Nanggulan Formation in Central Java (Lelono, 2000). The climate along the southern margin of Sundaland was very wet and supported extensive rain forests with widespread peat formation. During the late Eocene, the climate in southeast Sunda remained wet, still with widespread peat formation in low-lying areas. However, late Eocene pollen diversities are less, possibly reflecting a cooling climate, as suggested by the immigration of *Podocarpus* into the region, which probably dispersed at this time from India (Morley, 2011).

Oligocene drying, cooling and the expansion of Dipterocarpaceae

During the Oligocene, with cooler global climates, seasonal climates became widespread across the region, especially to the north of the paleo-equator, and peat formation became restricted to coastal areas south of Vietnam and the southern Sunda margin. The Oligocene has been most intensively studied palynologically in the pull-apart basins of the northern Sunda Shelf (Fig. 1, 30 Mya), where in the absence of marine index fossils, the stratigraphic succession is subdivided using a sequence biostratigraphic approach based on the recognition of climatostratigraphic cycles using palynology (Morley et al., 2003; Morley, 2011). These provide a powerful framework for stratigraphic correlation and understanding climate change (Fig. 3). Pollen floras differ markedly from those of the Eocene, with Poaceae and Pinaceae becoming prominent, reflecting on the one hand the much drier climate, and on the other the development of widespread horst and graben topography (i.e., raised and depressed fault blocks, respectively) with conifer pollen being derived from montane forests growing on higher altitude horsts, in areas possibly affected by water stress (Fig. 1, 30 Mya and 25 Mya). The conifers included the genera *Abies*, *Picea* and *Tsuga*, and these exhibit successive pollen maxima which probably reflect periods of cooler climate and freezing temperatures in the mountains at times of expansion of glaciers at high latitudes (Morley, 2012). Climates were wetter along the southern Sunda margin, but even to the south intermittent drier periods are recorded which include pollen comparable to Australian *Austrobuxus/Dissiliaria* (Euphorbiaceae) suggesting sclerophyllous vegetation (Lelono & Morley, 2011).

Dipterocarp pollen first appears in low numbers in the early Oligocene, but in the latest Oligocene, corresponding to the late Oligocene thermal maximum (Zachos et al., 2001), *Shorea* type pollen is very abundant, often occurring with common Poaceae pollen, suggesting the widespread development of monsoonal dipterocarp forest (Morley, 2012). The Oligocene also witnessed the dispersal into the area of *Dacrydium*, probably initially in lowland oligotrophic Kerangas or Kerapah settings where it is frequently associated with *Casuarina* type pollen in coals, suggesting ombrotrophic peat formation in a superwet climate (Morley, 2012). Kerapah swamps thus represent one of the oldest (and currently most threatened) ecosystems in the region.

Early Oligocene pollen floras were of markedly lower diversity than those from the Eocene, but during the late Oligocene, pollen diversities gradually increase, so that by the end of the Oligocene most of the pollen types characteristic of the wet, rain forest-dominated early Miocene are in place (Jais, 1997).

Early and middle Miocene pollen diversity and the evolution of the mangrove genus Sonneratia

The Early and middle Miocene are particularly noteworthy from the perspective of pollen diversity. Despite the seemingly optimal setting for taxon evolution with widespread rain forests, there is minimal change in pollen assemblage diversity for the period from about 21 Mya to 15 Mya (Fig. 3), possibly suggesting a very long period of floristic stability. On the other hand, the genus *Sonneratia* (Sonneratiaceae), in which all species can be differentiated on pollen characters (Muller, 1969) displays a remarkable diversification, with the sudden appearance of *Sonneratia caseolaris* pollen at 22 Mya and *S. alba* at 18 Mya. This reflects the successive adaptation of *Sonneratia* from freshwater swamp ancestors related to *Lagerstroemia* (Lythraceae) to more saline backmangrove and mangrove habitats (Graham et al., 2005). It is suggested that the extensive early Miocene brackish inland seas of the Sunda Shelf (Fig. 1, 25 Mya and 20 Mya) provided the niches for diversification of Sonneratiaceae and possibly other mangroves.

Extant Plant Diversity

Differences in geographical dynamics and constraints among major biogeographic regions through the last glacial cycle

The strong difference in floristic richness and, particularly, composition between west and east of the Wallace Line is likely strongly related to differences in climate change through the Quaternary Period. Several anecdotal examples illustrate the differences in taxonomic composition related to climatic differences, particularly as many evergreen everwet tree species appear to be intolerant of seasonal climates due to their geographic distribution. An obvious example is the almost complete lack of the Dipterocarpaceae in the eastern part of the SE Asian archipelago (Ashton, 1982), as these large emergent trees play a major role in ecosystem structure and dynamics and achieve remarkable levels of diversity on Sundaland while only a few species are found east of Wallace's Line. A similar pattern is true for the tropical Fagaceae in the genera *Lithocarpus* and *Castanopsis* (Soepadmo, 1972). While the region involves an increasing amount of contact between the ancient Laurasian and Gondwanan floras through time, previous paleo-climate reconstructions (Dam et al., 2001) suggest that the historical climate was substantially different on either side of Wallace's Line. As noted above, the eastern islands typically have lower levels of species diversity and local composition can be occasionally dominated by a few species. The forests of Sulawesi can be quite similar in structure and composition to the forests of Indochina, particularly in Myanmar and northern Thailand. A widespread species of *Castanopsis* forms almost mono-dominant stands at mid-elevations across the island of Sulawesi (Cannon et al., 2005). An almost identical species is found throughout mid-latitudes in Indochina (C. Cannon, pers. obs.), indicating a similar climate zone. A major question in relation to this difference in climate between west and east is to what degree dispersal limitation prevents taxa like the Dipterocarpaceae invading east, and how much of the limitation is historical climate change, where much of the east was too seasonal for true evergreen rainforest groups.

Finally, the species-rich peat swamps of SE Asia are a particularly unique feature of the region (Rieley & Page, 1995). Historical models suggest the conditions in the large central part of the Sunda Shelf during most of the Quaternary Period would have been favourable for the development of peat swamp deposits and that at certain points in sea-level change, the land area appropriate for the development of peat forests would have been tremendous, causing brief, thousand year periods of enormous population expansion. These forests persist in extreme environments on extremely poor soils, with trees needing to be tolerant to both drought and nutrient-poor conditions, and the floristic composition is composed of a long list of taxa endemic to these forests (Cannon & Leighton, 2004). More detailed study of the geographic distribution of genetic variation within and among the modern regions of peat swamp forests could be highly informative about the historical dynamics of rainforest through the Quaternary Period, primarily in the Sunda Shelf region, particularly population genetic studies to test whether the hypothesized burst in population occurred or not.

References

Ashton P.S. (1982) *Dipterocarpaceae.* Martinus Nijhoff, The Hague.

Barber A.J. & Crow M.J. (2003) An evaluation of plate tectonic models for the development of Sumatra. *Gondwana Research*, **6**, 1–28.

Barber A.J. & Crow M.J. (2009) Structure of Sumatra and its implications for the tectonic assembly of Southeast Asia and the destruction of Paleotethys. *Island Arc*, **18**, 3–20.

Boltenhagen E. (1976) Pollen et spores sénoniens du Gabon. *Cahiers de Micropaleontologie Nouvelle Serie*, **3**, 3–21.

Bosch D., Garrido C.J., Bruguier O., Dhuime B., Bodinier J.-L., Padròn-Navarta J.A. & Galland B. (2011) Building an island-arc crustal section: Time constraints from a LA-ICP-MS zircon study. *Earth and Planetary Science Letters*, **309**, 268–279.

Bouilhol P., Jagoutz O.E. & Hanchar J.M. (2010a) Isotopic constraints on the collision age from the Kohistan-Ladakh Arc crust. *American Geophysical Union, Fall Meeting 2010, unpublished abstract #T43B-2205*.

Bouilhol P., Schaltegger U., Chiaradia M., Ovtcharova M., Stracke A., Burg J.-P. & Dawood H. (2010b) Timing of juvenile arc crust formation and evolution in the Sapat Complex (Kohistan–Pakistan). *Chemical Geology*, **280**, 243–256.

Cannon C.H., Harting J.R. & Summers M. (2005) The vegetation of Sulawesi: II. fine filter analysis.

Cannon C.H. & Leighton M. (2004) Tree species distributions across five habitats in a Bornean rain forest. *Journal of Vegetation Science*, **15**, 257–266.

Dam R.A.C., van der Kaars S. & Kershaw A.P. (2001) Quaternary environmental change in the Indonesian region. *Palaeogeography, Palaeoclimatology, Palaeoecology*, **171**, 91–95.

Graham S.A., Hall J., Sytsma K. & Shi S.H. (2005) Phylogenetic analysis of the Lythraceae based on four gene regions and morphology. *International Journal of Plant Sciences*, **166**, 995–1017.

Hall R. (1996) Reconstructing Cenozoic SE Asia. *Tectonic evolution of Southeast Asia* (ed. by R. Hall and D.J. Blundell), pp. 153–184. Geological Society, London.

Hall R. (2002) Cenozoic geological and plate tectonic evolution of SE Asia and the SW Pacific: computer-based reconstructions, model and animations. *Journal of Asian Earth Sciences*, **20**, 353–431.

Hall R. (2011) Australia-SE Asia collision: plate tectonics and crustal flow. *The Southeast Asian Gateway: history and tectonics of Australia-Asia collision* (ed. by R. Hall, M.A. Cottam, and M.E.J. Wilson), pp. 75–109. The Geological Society of London, London.

Hall R., Clements B. & Smyth H.R. (2009) Sundaland: basement character, structure and plate tectonic development. *Proceedings Indonesian Petroleum Association*, **33th Annual**, 1–27.

Hay W.W. (1996) Tectonics and climate. *Geologische Rundschau*, **85**, 409–437.

Jais J.M. (1997) *Oligocene to Pliocene quantitative stratigraphic palynology of the southern Malay Basin, offshore Malaysia.* University of Sheffield.

Khan S.D., Walker D.J., Hall S.A., Burke K.C., Shah M.T. & Stockli L. (2009) Did the Kohistan-Ladakh island arc collide first with India? . *Geological Society of America Bulletin*, **121**, 366–384.

Kubitzki K. (2005) Major evolutionary advances in the history of green plants. *Acta Phytotaxonomica et Geobotanica*, **56**, 1–10.

Van Leeuwen T.M., Allen C.M., Kadarusman A., Elburg M., Palin J.M., Muhardjo & Suwijanto (2007) Petrologic, isotopic, and radiometric age constraints on the origin and tectonic history of the Malino Metamorphic Complex, NW Sulawesi, Indonesia. *Journal of Asian Earth Sciences*, **29**, 751–777.

Lelono E.B. (2000) *Palynology of the Nanggulan Formation, Central Java, Indonesia.* Royal Holloway, University of London.

Lelono E.B. & Morley R.J. (2011) Oligocene palynological succession from the East Java Sea. *Southeast Asian Gateway Evolution* (ed. by R. Hall and M. Wilson), pp. 333–345.

Metcalfe I. (1988) Origin and assembly of South-East Asian continental terranes. *Gondwana and Tethys* (ed. by M. Audley-Charles and A.G. Hallam), pp. 101–118. Oxford University Press, Oxford.
Metcalfe I. (1996) Gondwanaland dispersion, Asian accretion and evolution of Eastern Tethys. *Australian Journal of Earth Sciences*, **43**, 605–623.
Metcalfe I. (2005) South-East Asia. *Encyclopedia of geology* (ed. by R.C. Selley, L.R.M. Cocks, and I.R. Plimer), pp. 169–196. Elsevier, Amsterdam.
Metcalfe I. (2011a) Tectonic framework and Phanerozoic evolution of Sundaland. *Gondwana Research*, **19**, 3–21.
Metcalfe I. (2011b) Palaeozoic-Mesozoic history of SE Asia. *The Southeast Asian Gateway: history and tectonics of Australia-Asia collision* (ed. by R. Hall, M.A. Cottam, and M.E.J. Wilson), pp. 7–35. The Geological Society of London, London.
Metcalfe I. & Irving E. (1990) Allochthonous terrane processes in Southeast Asia. *Philosophical Transactions of the Royal Society of London A*, **331**, 625–640.
Morley R.J. (1998) Palynological evidence for Tertiary plant dispersals in the SE Asian region in relation to plate tectonics and climate. *Biogeography and Geological Evolution of SE Asia* (ed. by R. Hall and J.D. Holloway), pp. 211–234. Backhuys Publishers, Leiden.
Morley R.J. (2000) *Origin and evolution of tropical rain forests.* John Wiley & Sons, Chichester.
Morley R.J. (2001) Why are there so many primitive angiosperms in the rainforests of Asia-Australia? *Floral and faunal migrations and evolution in SE Asia-Australia* (ed. by I. Metcalfe, J.M.B. Smith, M. Morwood, and I. Davidson), pp. 185–200. Lisse.
Morley R.J. (2006) Cretaceous and Tertiary climate change and the past distribution of megathermal rain forests. *Tropical rainforest responses to climatic change* (ed. by M. Bush and F. J.R.), pp. 1–54. Springer Praxis Books, Berlin.
Morley R.J. (2011) Dispersal and paleoecology of tropical podocarps. *Ecology of the Podocarpaceae in tropical forests* (ed. by B.L. Turner and L.M. Cernusak), pp. 21–41. Smithsonian Institution Scholarly Press, Washington, D.C.
Morley R.J. (2012) A review of the Cenozoic palaeoclimate history of Southeast Asia. *Biotic evolution and environmental change in Southeast Asia* (ed. by D. Gower, K.G. Johnson, B.R. Rosen, J. Richardson, L. Rüber, and S.T. Williams), pp. 79–114. Cambridge University Press, Cambridge.
Morley R.J., Morley H.P. & Restrepo-Pace P. (2003) Unraveling the tectonically controlled stratigraphy of the West Natuna Basin by means of palaeo-derived Mid Tertiary climate changes. *Proceedings Indonesian Petroleum Association*, **29th Annual**, 1–24.
Muller J. (1968) Palynology of the Pedawan and Plateau Sandstone Formations (Cretaceous-Eocene) in Sarawak. *Micropaleontology*, **14**, 1–37.
Muller J. (1969) A palynological study of the genus Sonneratia (Sonneratiaceae). *Pollen et Spores*, **11**, 223–298.
Rieley J.O. & Page S.E. (1995) *Biodiversity and sustainability of tropical peatlands* . Samara Publishing Limited, Cardigan, UK.
Smyth H.R., Hamilton P.J., Hall R. & Kinny P.D. (2007) The deep crust beneath island arcs: Inherited zircons reveal a Gondwana continental fragment beneath East Java, Indonesia. *Earth and Planetary Science Letters*, **258**, 269–282.
Soepadmo E. (1972) *Fagaceae.* Noordhoff International Publishing, Leyden.
Sone M. & Metcalfe I. (2008) Parallel Tethyan sutures in mainland SE Asia: New insights for Palaeo-Tethys closure and implications for the Indosinian orogeny. *Comptes Rendus Geoscience*, **340**, 166–179.
Sun G., Dilcher D.L., Zheng S. & Zhou Z. (1998) In search of the first flower: a Jurassic angiosperm, Archaefructus, from northeast China. *Science*, **282**, 1692–1695.
Swisher III C.C., Wang X., Zhou Z., Wang Y., Jin F., Zhang J., Xu X. & Zhang F. (2002) Further support for a Cretaceous age for the feathered-dinosaur beds of Liaoning, China: New 40Ar/39Ar dating of the Yixian and Tuchengzi Formations. *Chinese Science Bulletin*, **47**, 136–139.
Takhtajan A. (1969) *Flowering plants: origin and dispersal (translated by C. Jeffrey)*. Oliver & Boyd, Edinburgh.

Wakita K. & Metcalfe I. (2005) Ocean plate stratigraphy in East and Southeast Asia. *Journal of Asian Earth Sciences*, **24**, 679–702.

Zachos J., Pagani M., Sloan L., Thomas E. & Billups K. (2001) Trends, rhythms, and aberrations in global climate 65 Ma to present. *Science*, **292**, 393–686.

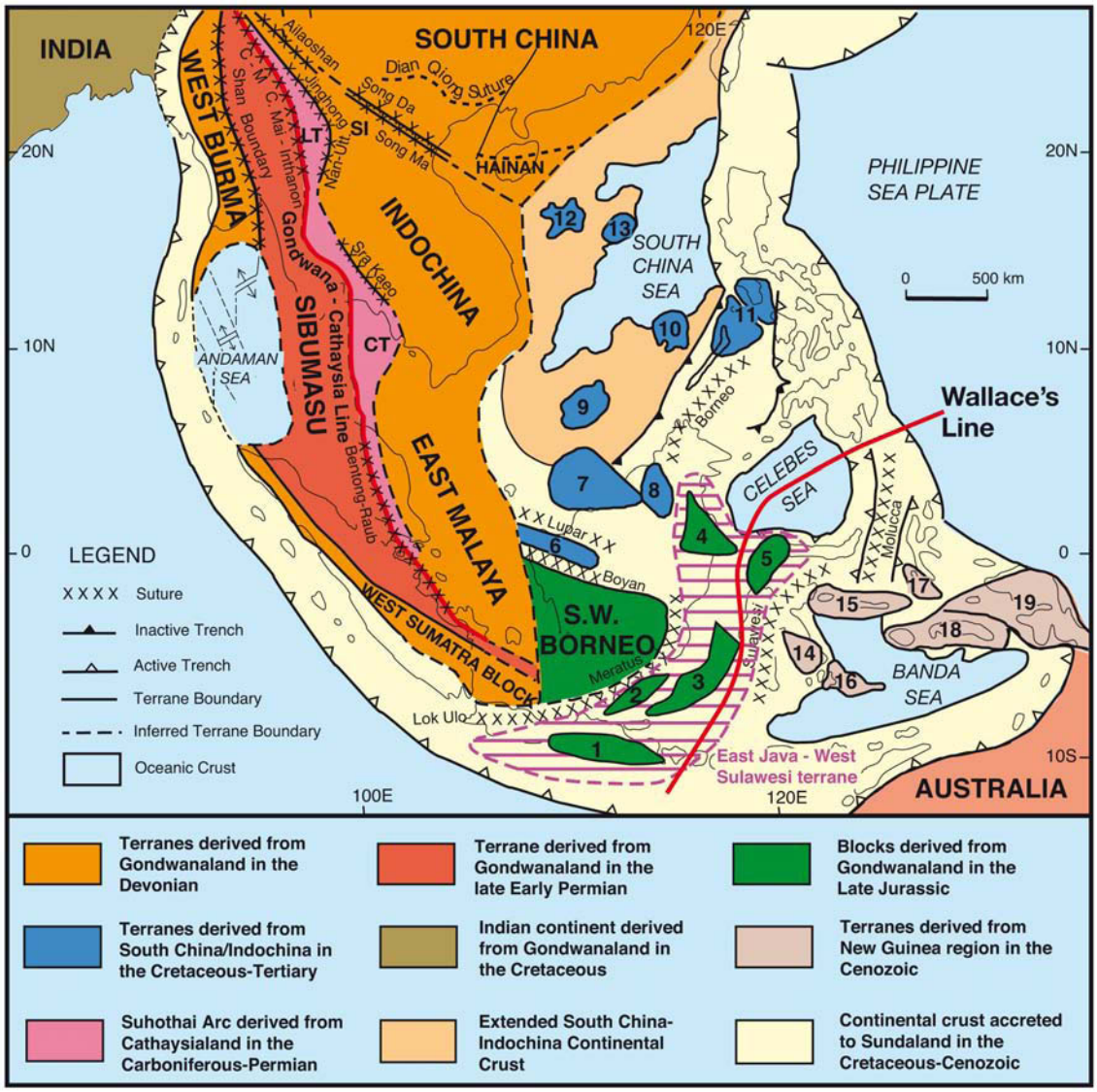


Figure S1. Distribution of continental blocks, fragments and terranes, and principal sutures of SE Asia. Numbered microcontinental blocks: 1, East Java; 2, Bawean; 3, Paternoster; 4, Mangkalihat; 5, West Sulawesi; 6, Semitau; 7, Luconia; 8, Kelabit-Longbowan; 9, Spratly Islands-Dangerous Ground; 10, Reed Bank; 11, North Palawan; 12, Paracel Islands; 13, Macclesfield Bank; 14, East Sulawesi; 15, Bangai-Sula; 16, Buton; 17, Obi-Bacan; 18, Buru-Seram; 19, West Irian Jaya. LT, Lincang Terrane; CT, Chanthaburi Terrane; C-M, Changning-Menglian Suture; C.-Mai – Inthanon, Chiang Mai – Inthanon Suture; Nan-Utt., Nan-Uttaradit Suture. Hatched area shows interpreted East Java – West Sulawesi terrane. The Paleozoic Gondwana - Cathaysia and extant Wallace biogeographic lines are shown in red. Figure modified after Metcalfe (2011b).

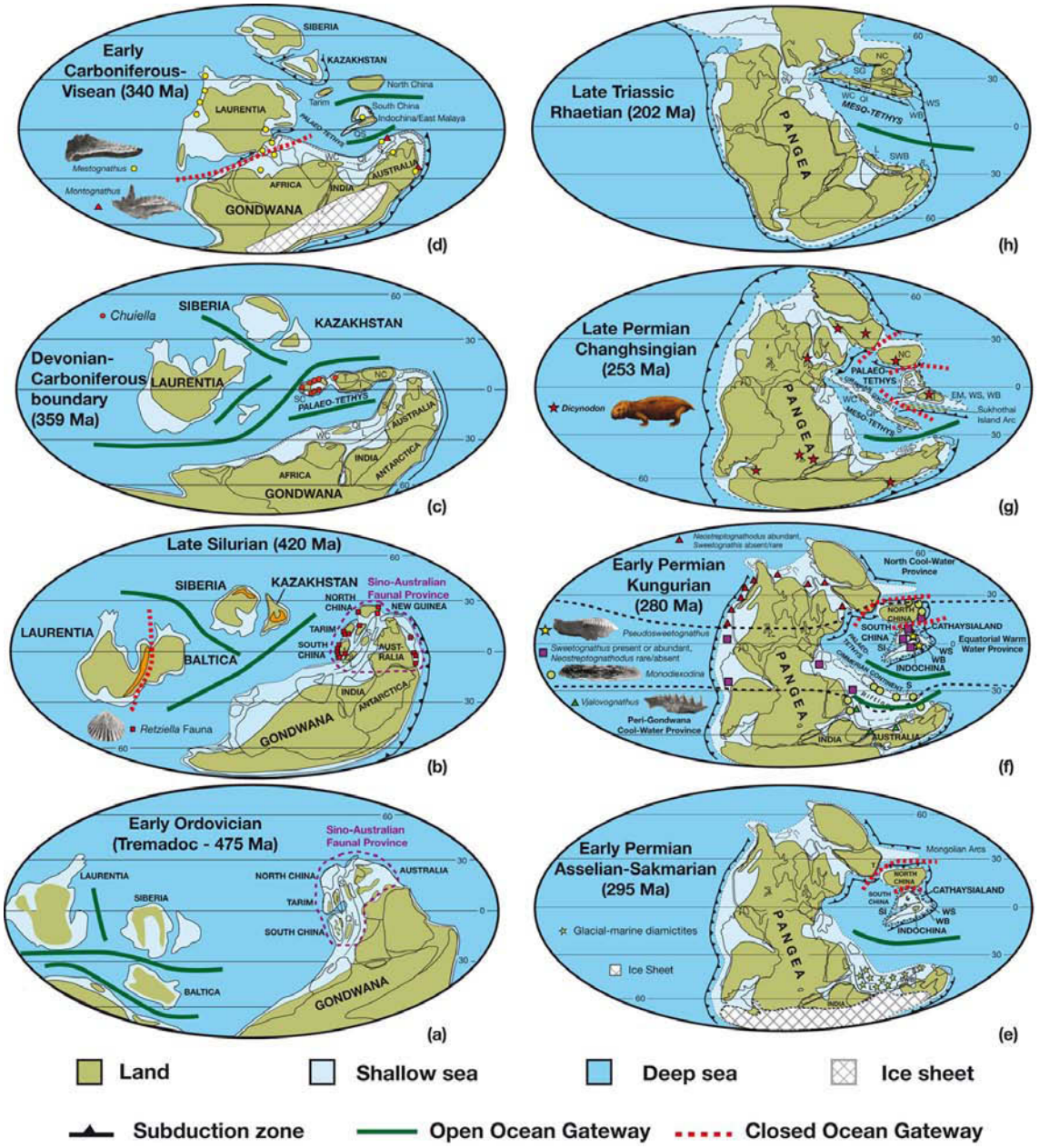


Figure S2. Global paleogeographic reconstructions illustrating the Paleozoic evolution of the Tethyan region from Ordovician to Triassic times showing relative positions of the Southeast Asian terranes, distribution of land and sea, evolution of ocean gateways, and some key biogeographic distributions. NC, North China; SG, Songpan Ganzi; SC, South China; WC, Western Cimmerian Continent; QI, Qiangtang Block; I, Indochina Block; SI, Simao Block; S, Sibumasu Block; EM, East Malaya Block; WS, West Sumatra Block; WB, West Burma Block; L, Lhasa Block; SWB, Argoland/SW Borneo.

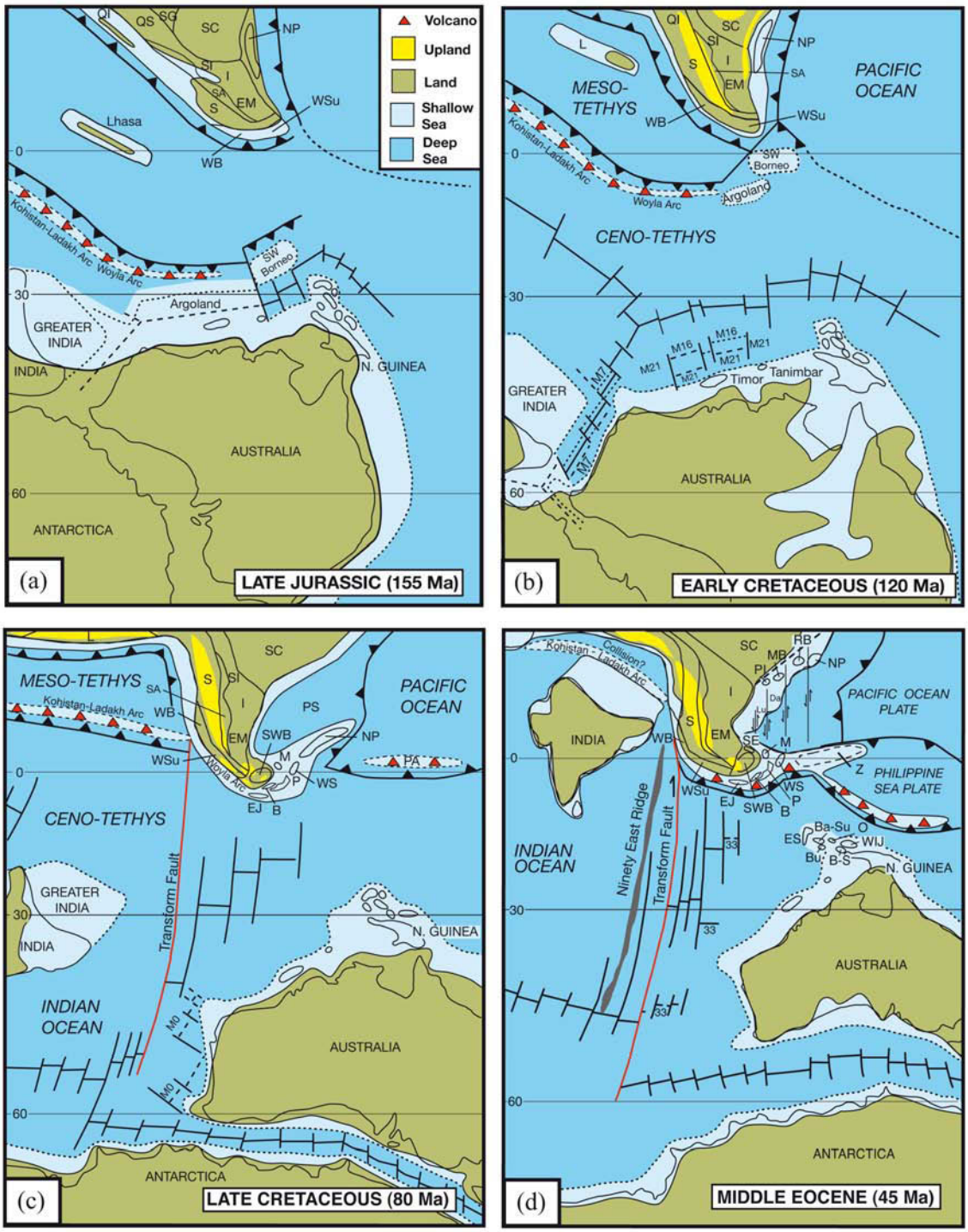


Figure S3. Paleogeographic reconstructions for Eastern Tethys in (a) Late Jurassic, (b) Early Cretaceous (c) Late Cretaceous and (d) middle Eocene showing distribution of continental blocks and fragments of SE Asia–Australasia and land and sea. Upland areas are tentative and do not necessarily represent high mountains. QS, Qamdo-Simao; SA, Sukhothai Arc; NP, North Palawan and other small continental fragments now forming part of the Philippines basement; M, Mangkalihat; WS, West Sulawesi; P, Paternoster; B, Bawean; PA, Incipient East Philippine Arc; PS, Proto-South China Sea; Z, Zambales Ophiolite; ES, East Sulawesi; O, Obi-Bacan; Ba-Su, Banggai-Sula; Bu, Buton; WIJ, West Irian Jaya. M numbers represent Indian Ocean magnetic anomalies. Other abbreviations as for Fig. S2. After Metcalfe (2011b).

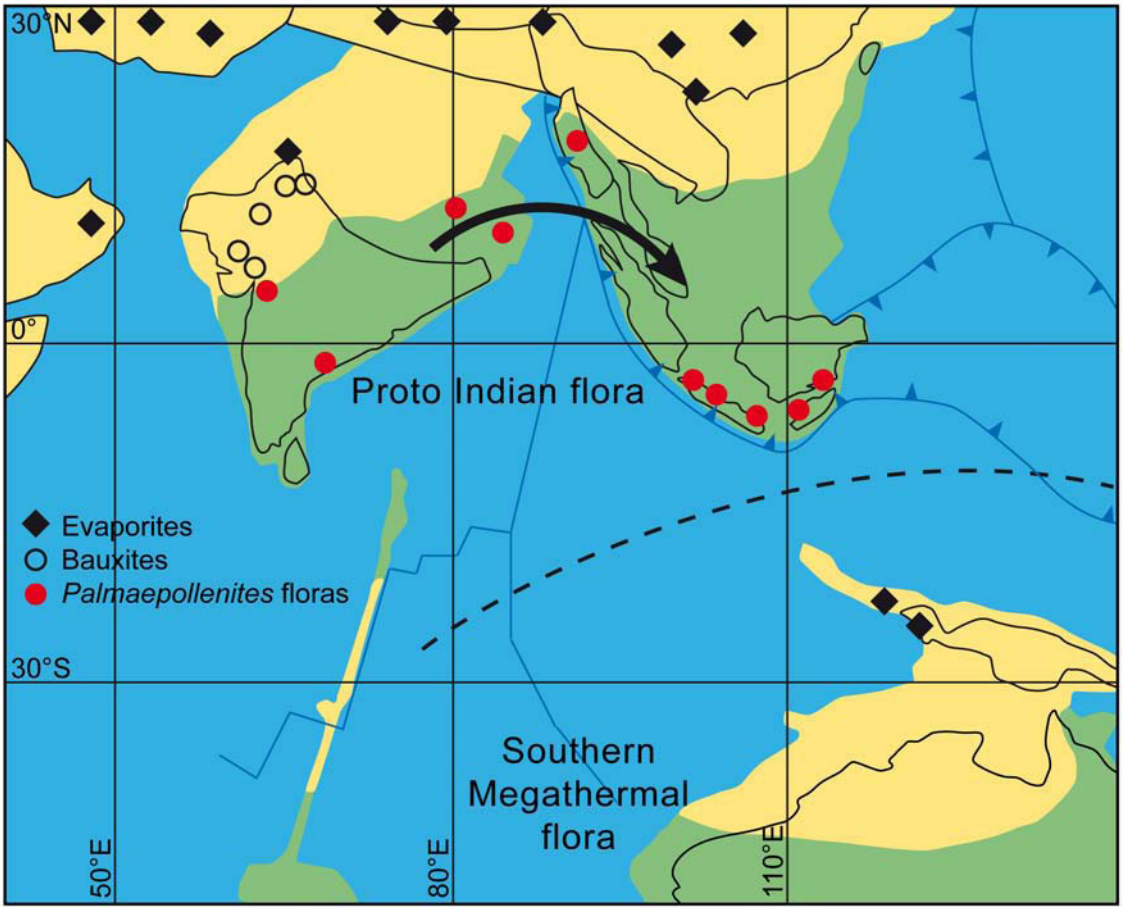


Figure S4. Middle Eocene (45 Mya) plate tectonic setting and paleogeography for India and Southeast Asia, showing the manner in which the northern drift of the Indian Plate brought the Indian and SE Asian floras into close juxtaposition, and within the same climatic zone, allowing elements of the more aggressive Indian flora to disperse widely into SE Asia, resulting in the extinction of many elements of the previous Paleocene flora in the Sunda region. The collision of the Indian and Asian plates is thought to have occurred from 50 Mya, but the extents of greater India and greater Asia are not agreed.

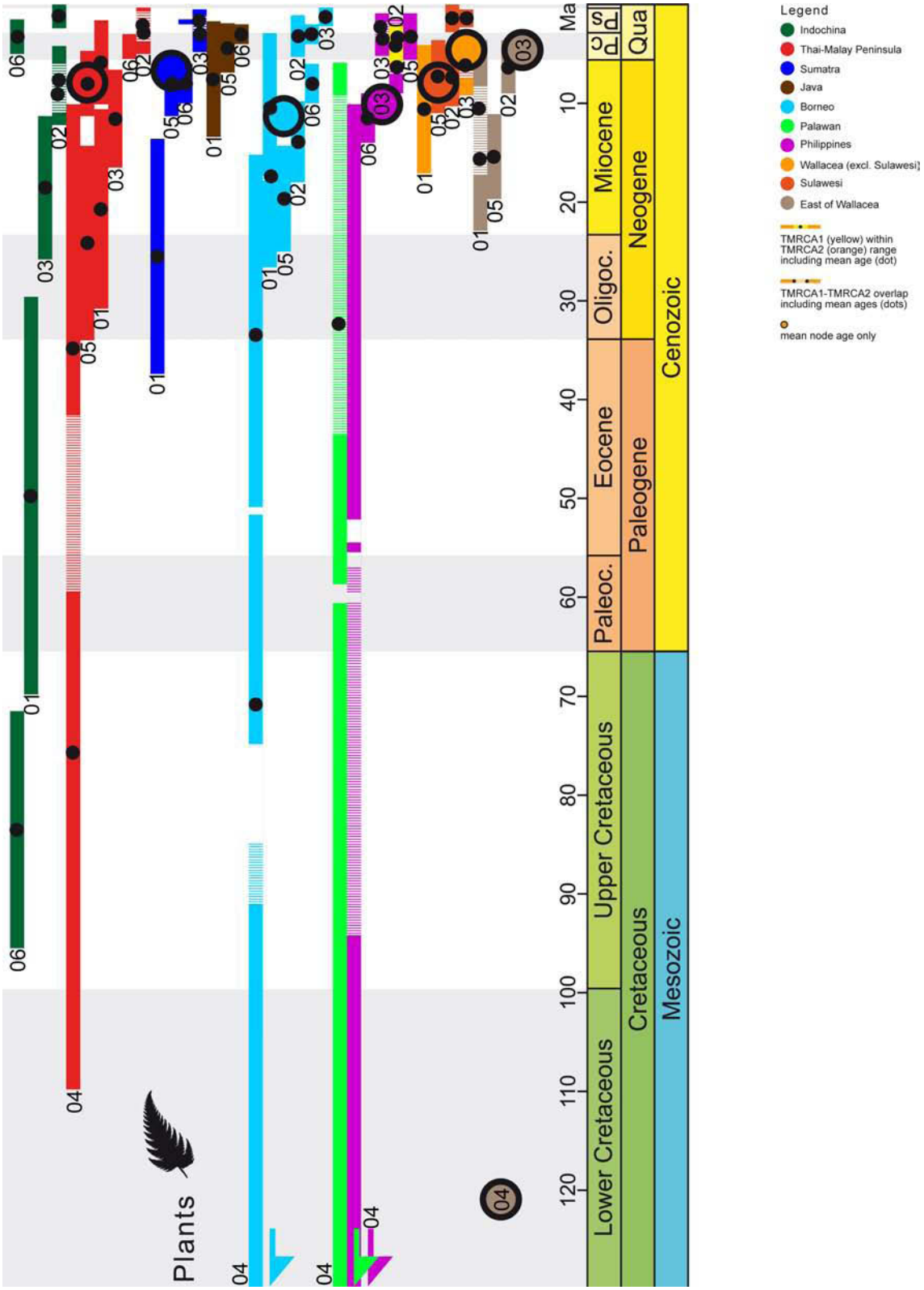


Figure S5. Distribution of relaxed-clock (BEAST) divergence times for plants. Numbers next to time bars correspond to the dataset numbers used in Table S1-5. Divergence events for which node support is <50 (and thus lacking time bars) are depicted as single color-coded circles (mean ages). 01: Meliaceae (*Aglaia*), 02: Araceae (*Alocasia*), 03: Begoniaceae (*Begonia*), 04: Gesneriaceae (*Cyrtandra*), 05: Ericaceae (*Rhododendron*), 06: Rafflesiaceae (several).

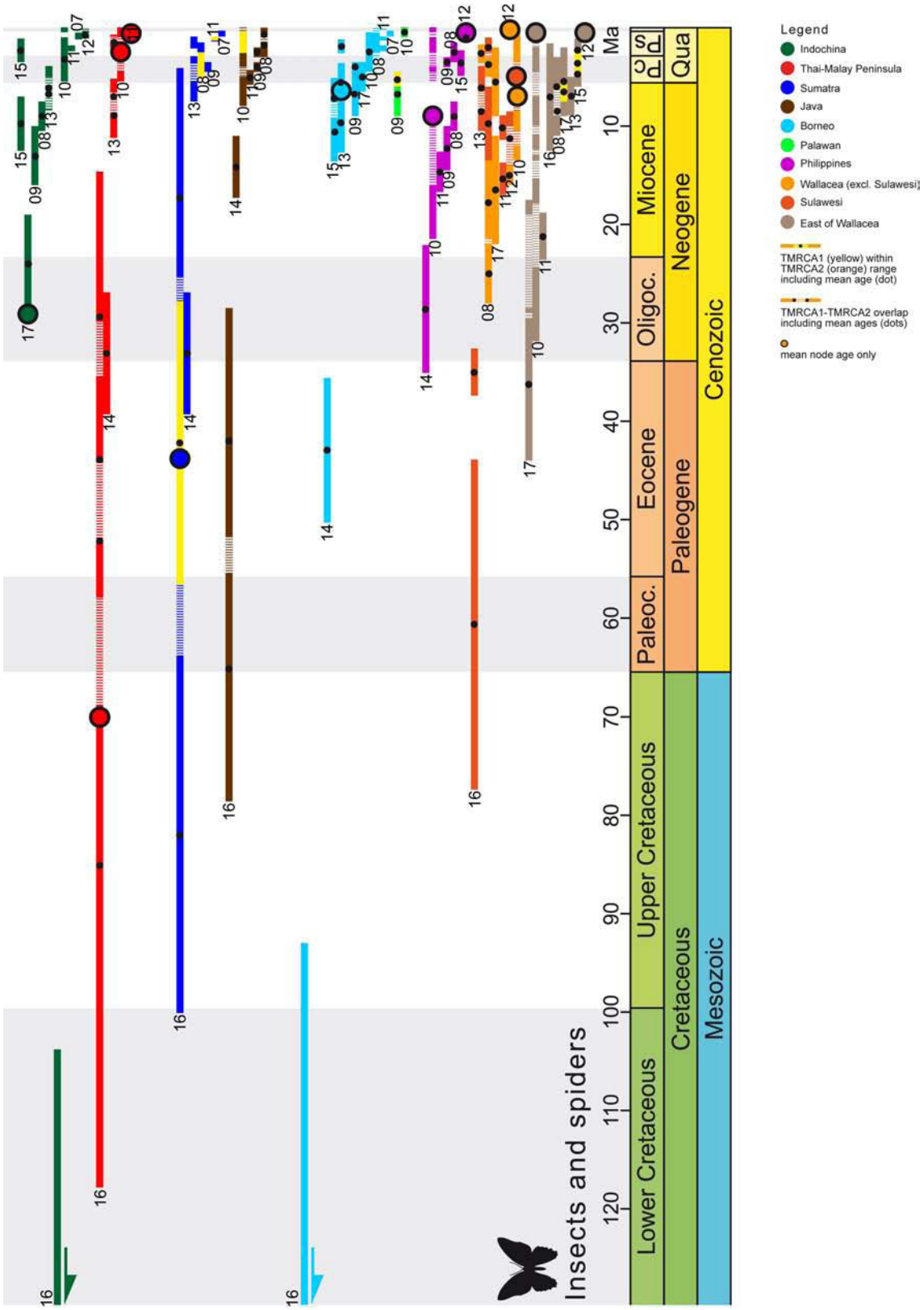


Figure S6. Distribution of relaxed-clock (BEAST) divergence times for insects and spiders. Numbers next to time bars correspond to the dataset numbers used in Table S1-4. 07: Culicidae (*Anopheles*), 08: Nymphalidae (*Cethosia*), 09: Nymphalidae (*Charaxes*), 10: Pieridae (*Delias*), 11: Lycidae (*Metriorrhynchus*), 12: Tetragnathidae (*Nephila*), 13: Blaberidae (*Salganea*), 14: Lycidae (*Scarelus*), 15: Blaberidae (several), 16: Stylocellidae (several), 17: Sycophaginae (several).

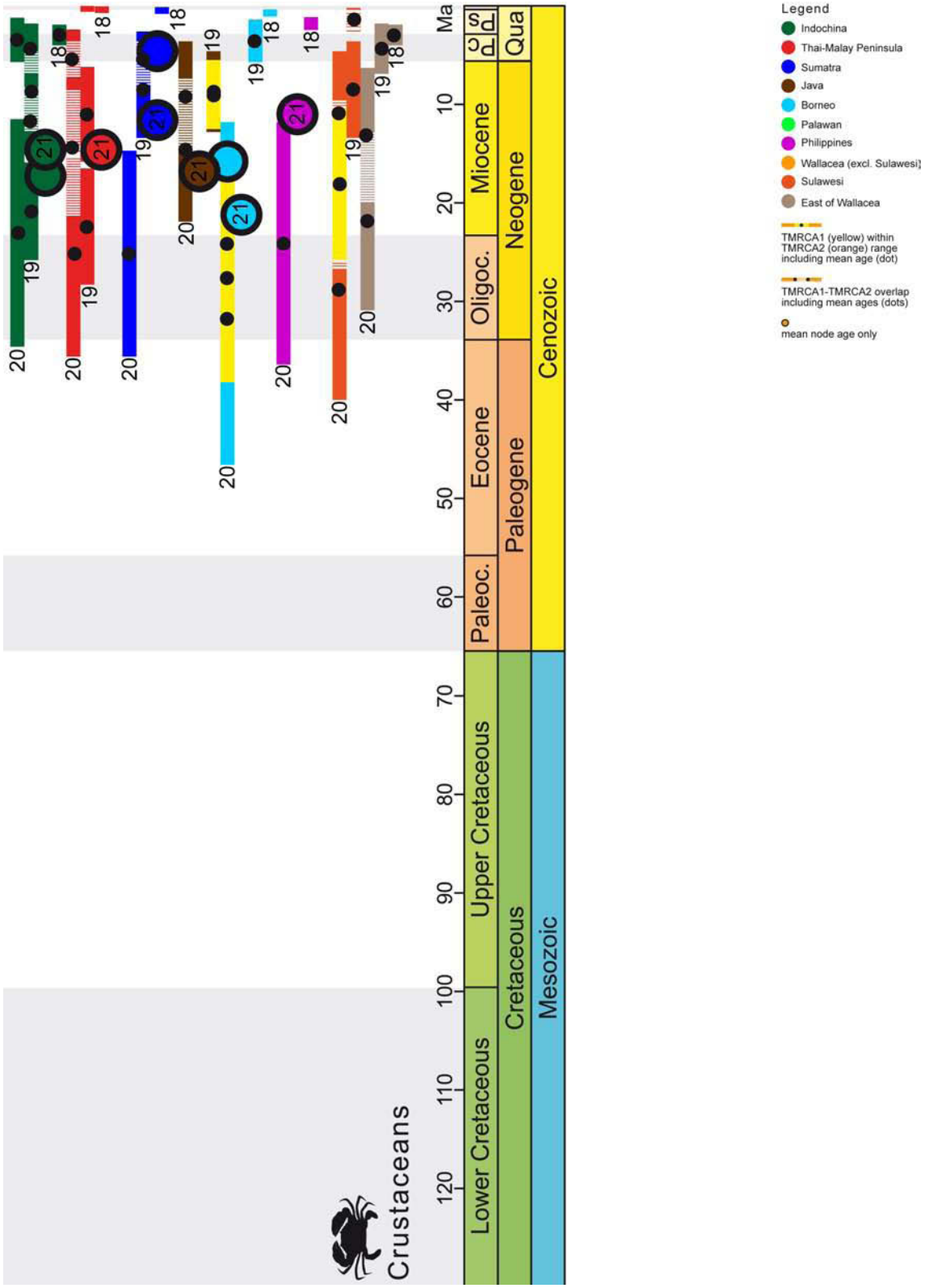


Figure S7. Distribution of relaxed-clock (BEAST) divergence times for freshwater crustaceans. Numbers next to time bars correspond to the dataset numbers used in Table S1-4. 18: Palaemonidae (*Macrobrachium*), 19: Palaemonidae (*Macrobrachium*), 20: Gecarcinucidae (several), 21: Potamidae (several).

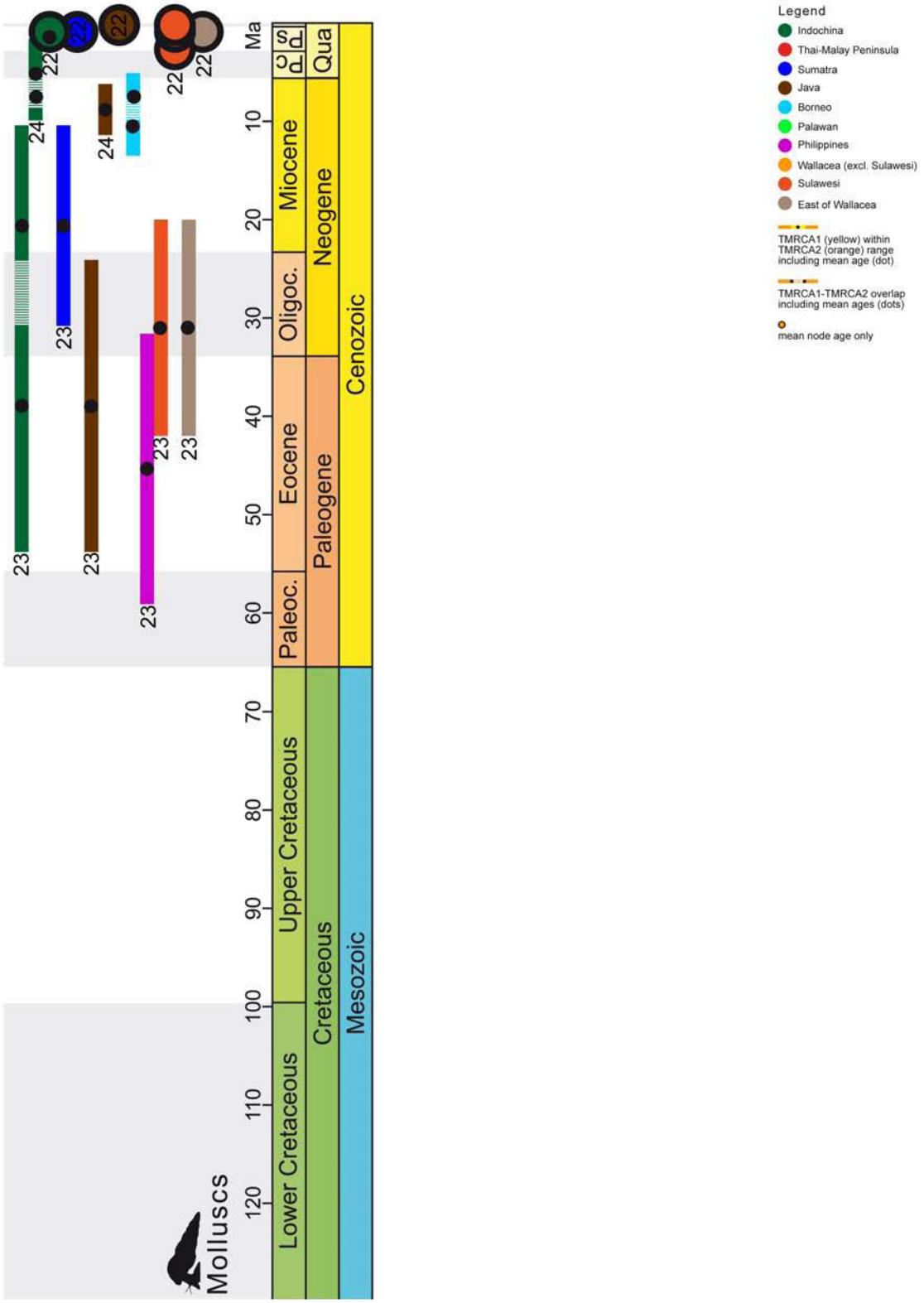


Figure S8. Distribution of relaxed-clock (BEAST) divergence times for freshwater molluscs. Numbers next to time bars correspond to the dataset numbers used in Table S1-4. 22: Corbiculidae (*Corbicula*), 23: Pachychilidae (several), 24: Pachychilidae (*Sulcospira*).

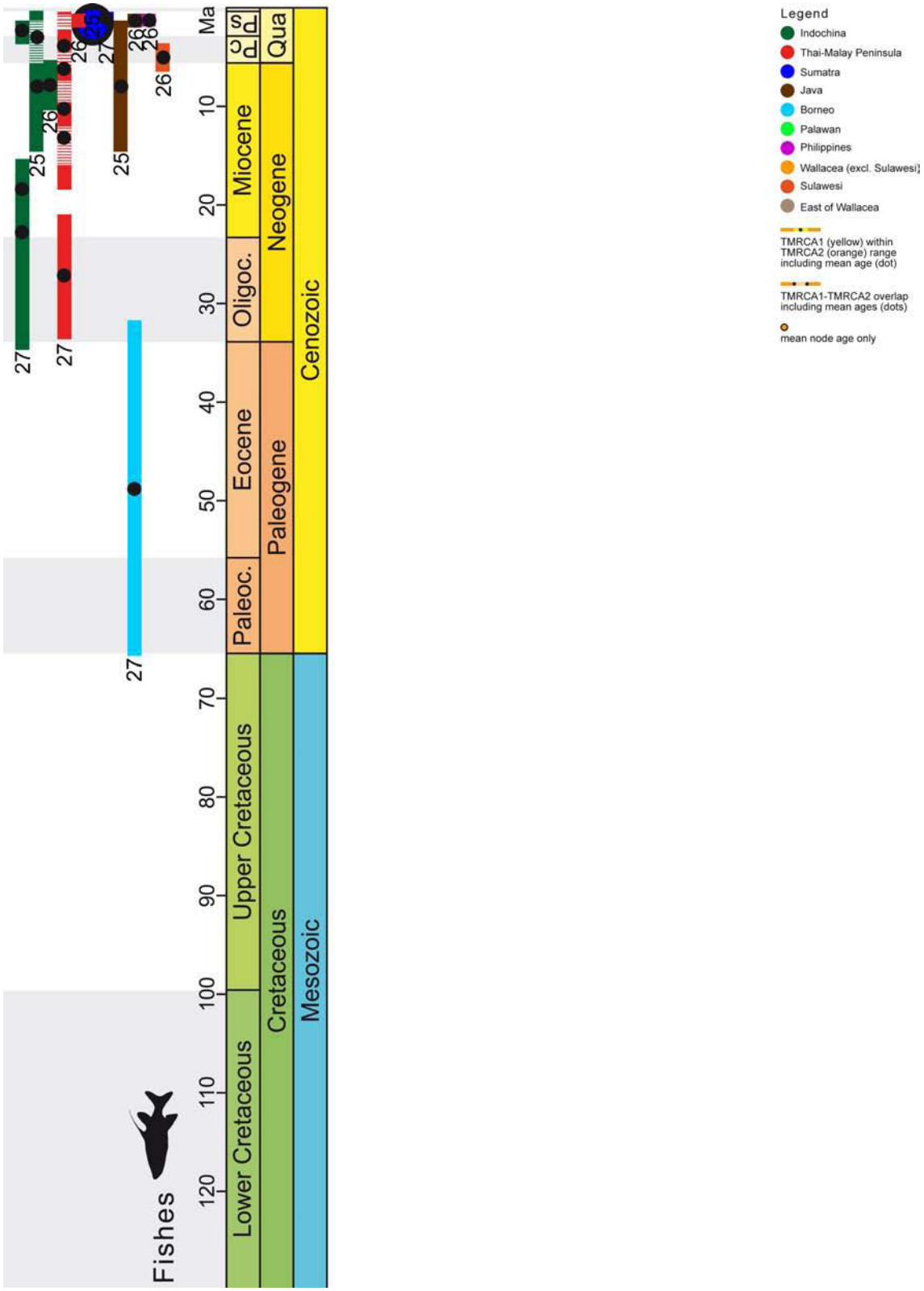


Figure S9. Distribution of relaxed-clock (BEAST) divergence times for freshwater fishes. Numbers next to time bars correspond to the dataset numbers used in Table S1-4. 25: Cyprinidae (*Barbodes*), 26: Adrianichthyidae (*Oryzias*), 27: Cobitidae (*Pangio*).

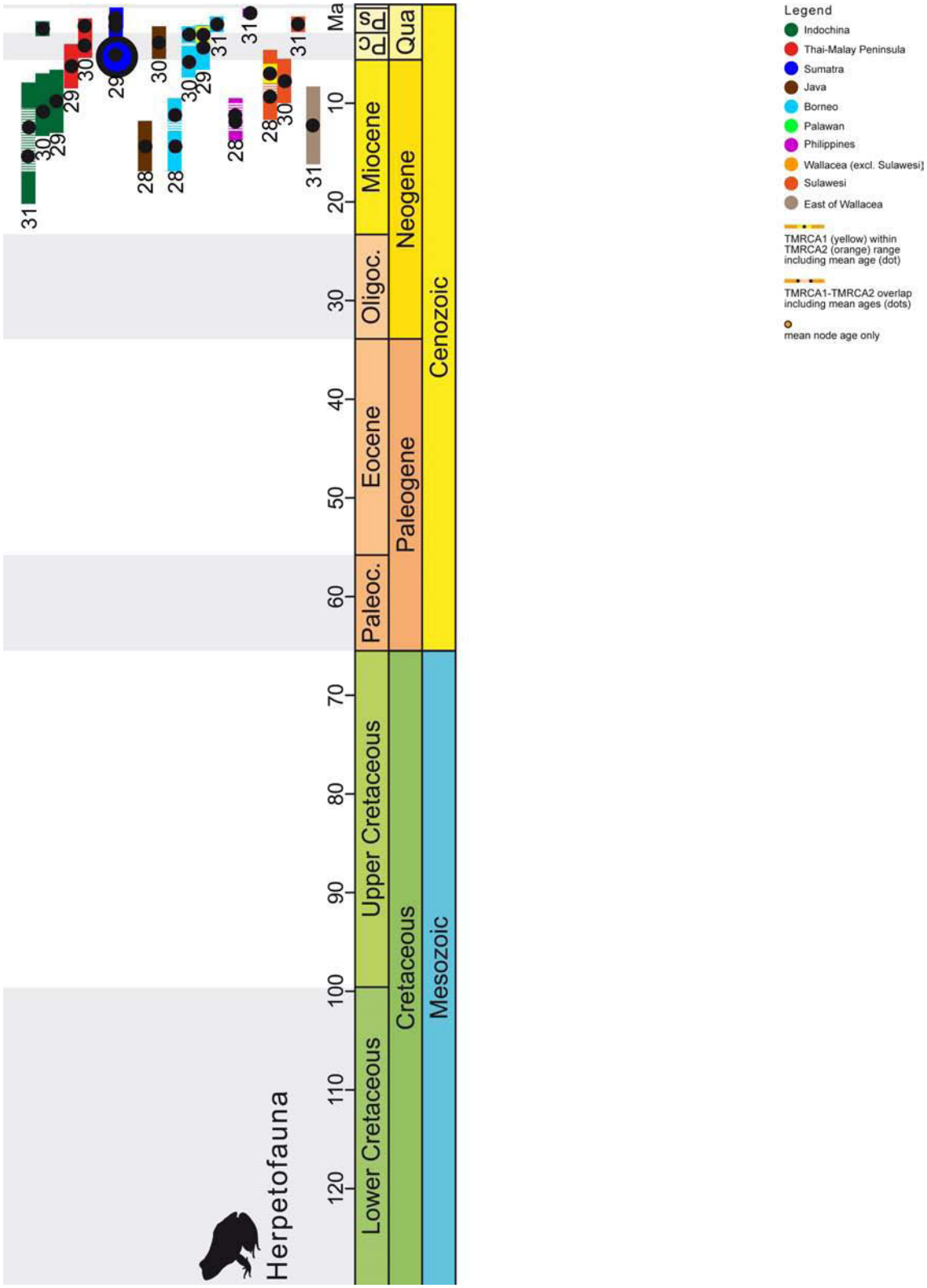


Figure S10. Distribution of relaxed-clock (BEAST) divergence times for amphibians and reptiles (herpetofauna). Numbers next to time bars correspond to the dataset numbers used in Table S1-4. 28: Dicroglossidae (*Limnonectes*), 29: Ranidae (*Rana*), 30: Ranidae (*Rana*), 31: Homalopsidae (several).

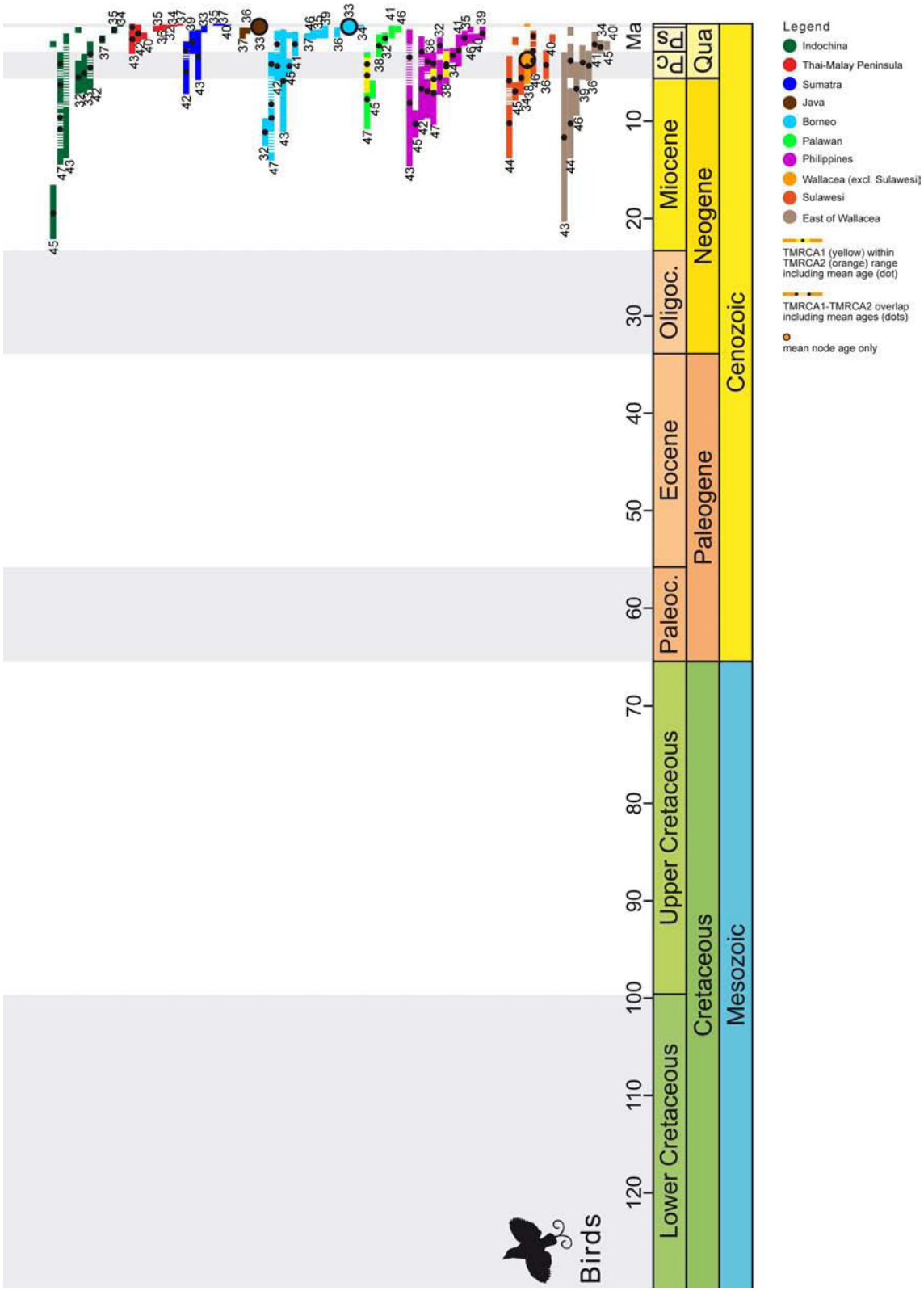


Figure S11. Distribution of relaxed-clock (BEAST) divergence times for birds. Numbers next to time bars correspond to the dataset numbers used in Table S1-4. 32: Nectariniidae (*Arachnothera*), 33: Cettiidae (*Cettia*), 34: Alcedinidae (*Ceyx*), 35: Turdidae (*Copsychus*), 36: Dicaeidae (*Dicaeum*), 37: Muscicapidae (*Enicurus*), 38: Muscicapidae (*Ficedula*), 39: Campephagidae (*Lalage*), 40: Nectariniidae (*Nectarinia*), 41: Phylloscopidae (*Phylloscopus*), 42: Pycnonotidae (*Pycnonotus*), 43: Rhipiduridae (*Rhipidura*), 44: Megapodiidae (several), 45: Zosteropidae (several), 46: Pachycephalidae (several), 47: Pycnonotidae (several).

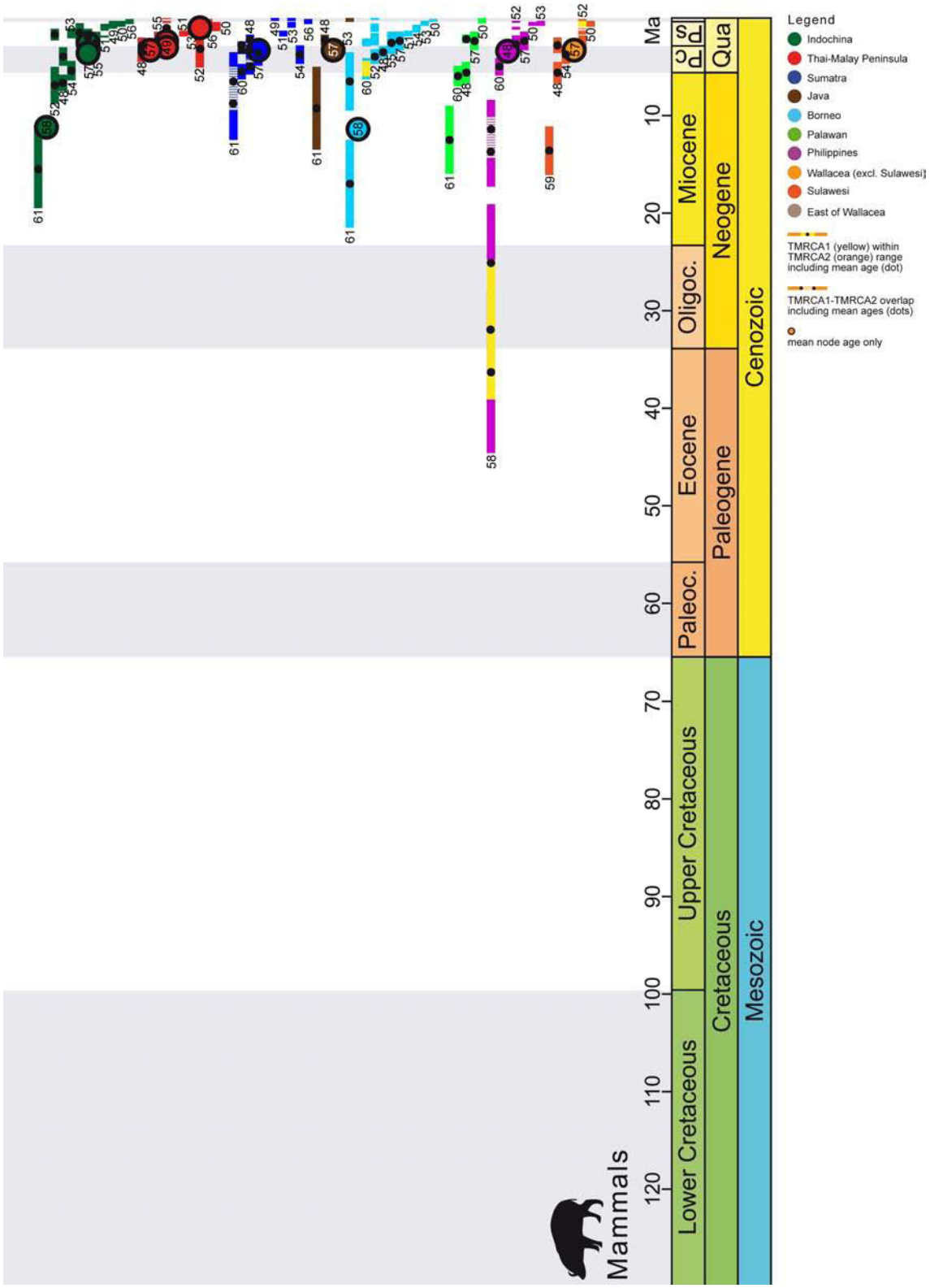


Figure S12. Distribution of relaxed-clock (BEAST) divergence times for mammals. Numbers next to time bars correspond to the dataset numbers used in Table S1-4. 48: Soricidae (*Crocidura*), 49: Canidae (*Cuon*), 50: Pteropdidae (*Cynopterus*), 51: Elaphantidae (*Elaphas*), 52: Hipposideridae (*Hipposideros*), 53: Cercopithecidae (*Macaca*), 54: Cercopithecidae (*Macaca*), 55: Felidae (*Neofelis*), 56: Felidae (*Panthera*), 57: Viverridae (*Paradoxurus*), 58: Muridae (several), 59: Sciuridae (several), 60: Sciuridae (*Sundascirus*), 61: Tupaiidae (*Tupaia*).

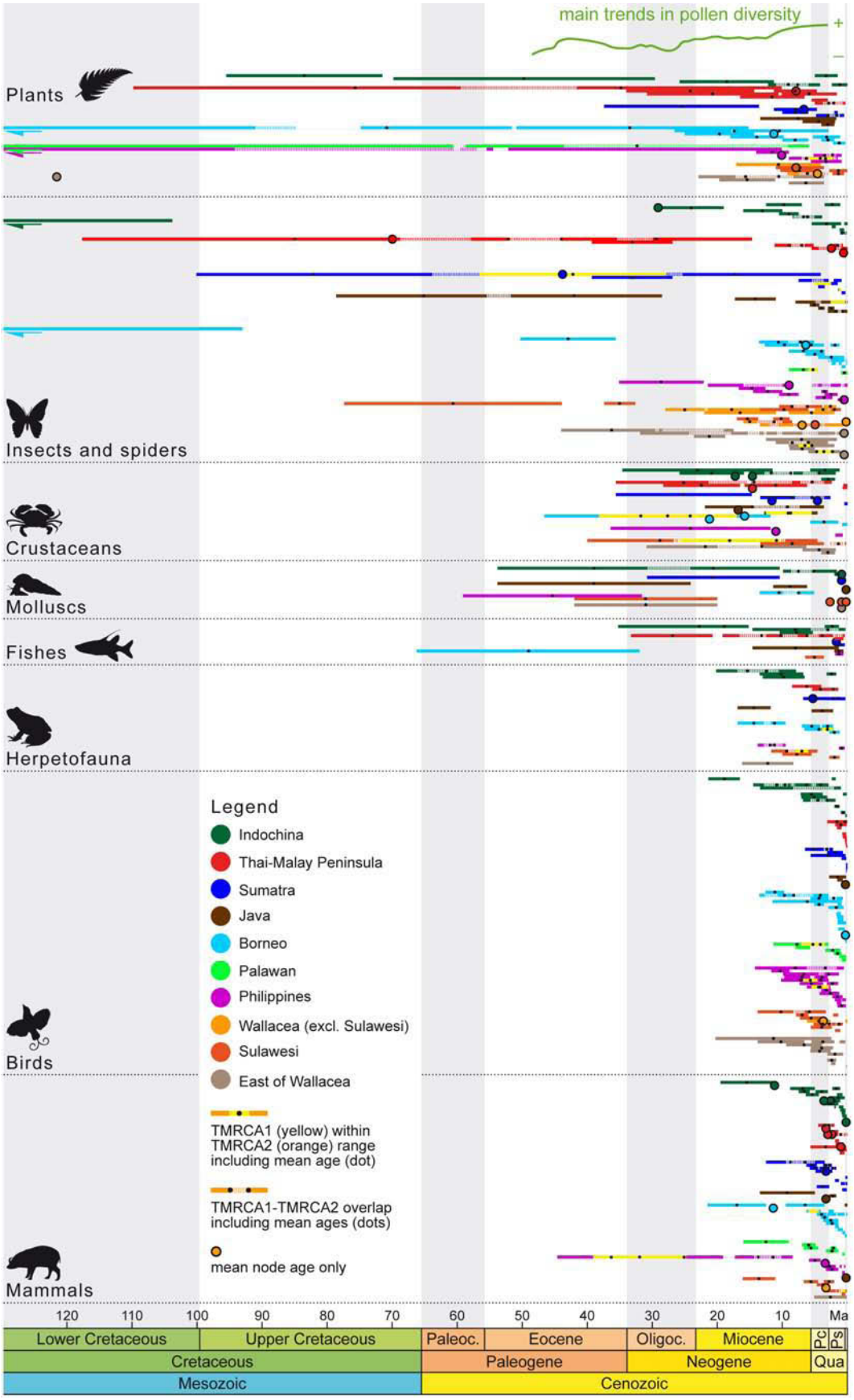


Figure S13. Distribution of divergence times for all study clades. Bars represent 95% highest posterior density (HPD) and mean of estimated node ages. Within each major group (plants, fishes, etc.), nodes for all phylogenies are lumped together and then sorted vertically by area. Large dots represent nodes in MCC trees with <0.5 posterior support. Multiple bars in the same plane denote multiple monophyletic area-clades identified (see Material and Methods for details).

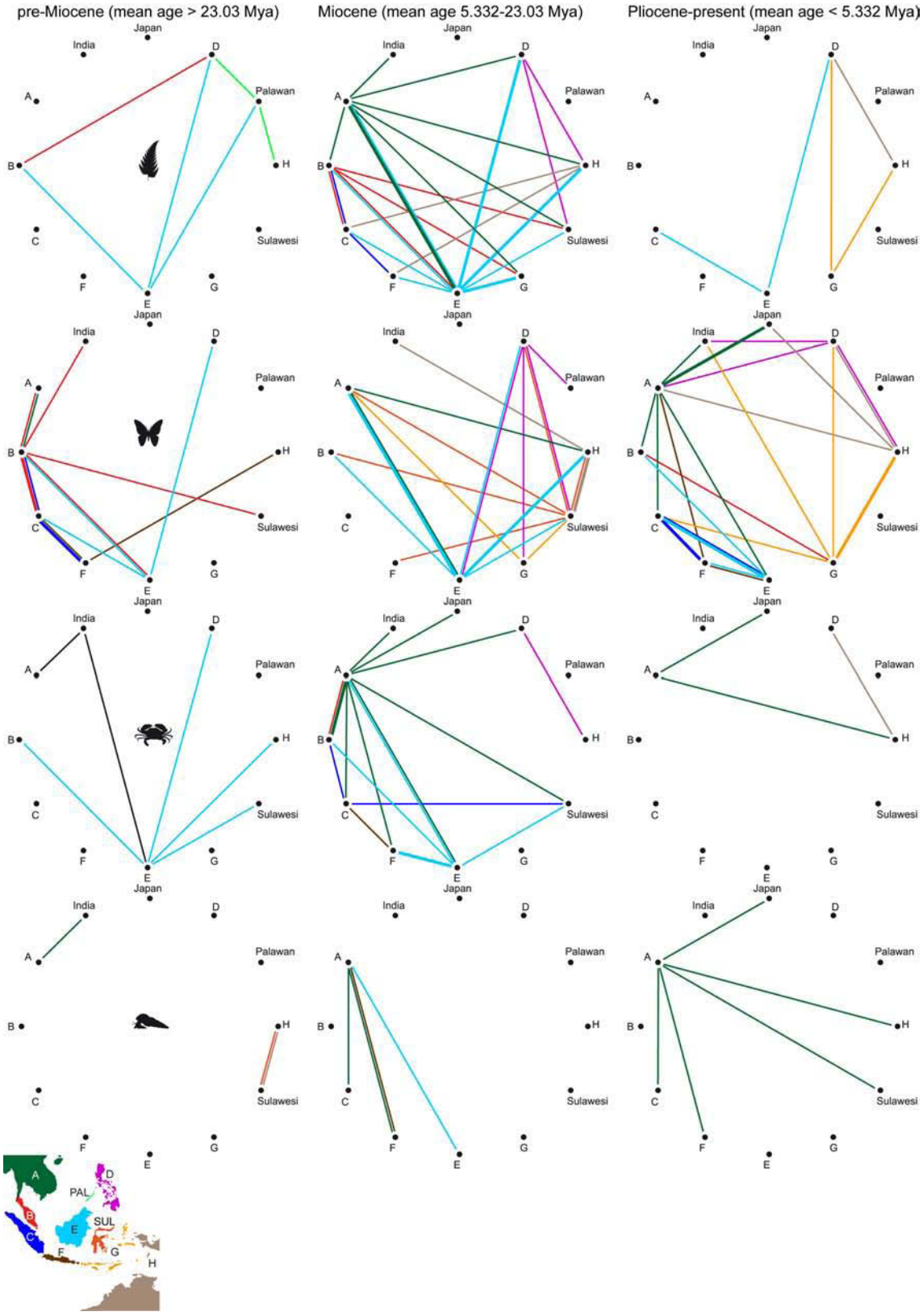


Figure S14. Total colonization routes through time (pre-Miocene, Miocene, Plio-Pleistocene) inferred from ancestral area reconstructions for plants, insects and spiders, freshwater crustaceans, and freshwater molluscs.

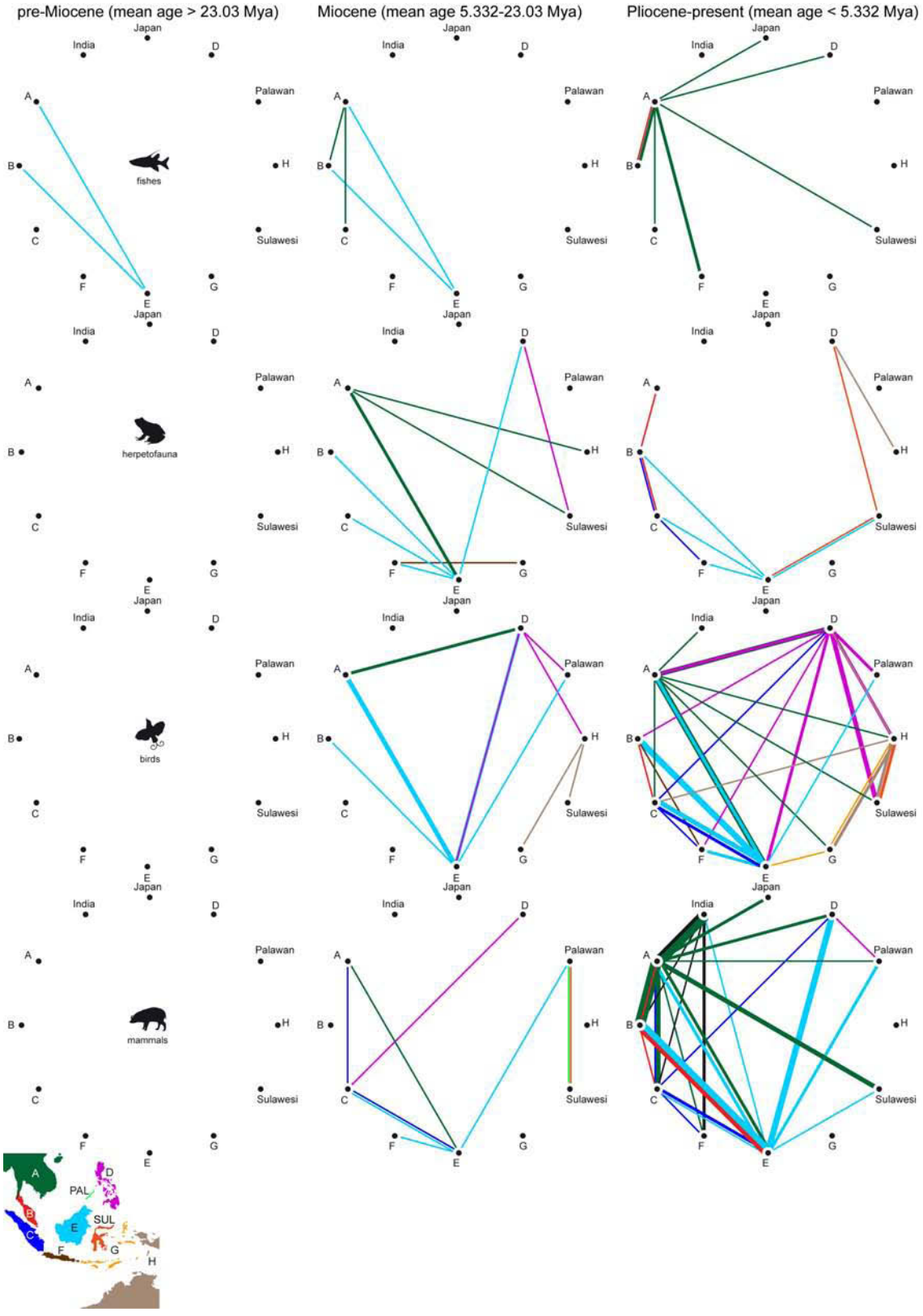


Figure S15. Total colonization routes through time (pre-Miocene, Miocene, Plio-Pleistocene) inferred from ancestral area reconstructions for freshwater fishes, amphibians and reptiles (herpetofauna), birds, and mammals.

Table S1 (next pages). List of studied datasets including inferred min/max ages (95% HPD = credibility intervals) of the time to the most recent common ancestor (TMRCA) for each lineage/monophyletic group corresponding to the defined areas inferred from molecular clock analyses. Multiple colonization events are denoted by vertical lines. Divergence times from Sulawesi are included in the 'TMRCA G' column (Wallacea), but marked red.

TMRCA 95% HPD (credibility intervals)		TMRCA IND India/Sri Lanka	India/Sri Lanka	TMRCA A Indochina	Indochina	TMRCA B Malay peninsula	Malay peninsula
dataset	genus	min	max	min	max	min	max
01	*Aglaia*	21 5	21 5	29 6	69 8	10 6 | 1 6	30 8 | 10 1
02	*Alocasia*			4 2 | 6 0 | 0 0	11 0 | 12 2 | 2 4	0 3 | 0 7	3 8 | 5 0
03	*Begonia*			11 3	25 8	6 6	16 5
04	*Cyrtandra*					41 6 | 10 1	109 8 | 59 5
05	*Rhododendron*					14 3 | 7 8 | 4 7	24 0 | 7 8 | 11 3
06	several			71 5 | 1 5	95 5 | 5 0	3 0	5 5
07	*Anopheles*			0 5	1 4	0 3	0 8
08	*Cethosia*	16 5	26 5	7 5	10 5		
09	*Charaxes*	3 5	4 5	10 0	16 0		
10	*Delias*	4 0	10 5	0 0 | 1 0	0 5 | 5 5	0 0 | 0 0 | 0 0 | 0 5 | 1 0 | 2 5	0 5 | 2 5 | 3 0 | 4 5 | 5 5 | 2 5
11	*Metriorrhynchus*			1 8	2 4	0 5 | 0 5	0 5 | 0 5
12	*Nephila*	0 1 | 0 2	0 5 | 0 7	0 1 | 0 3 | 0 1	0 8 | 0 9 | 0 5	0 3	0 7
13	*Salganea*	6 6	12 7	3 9 | 5 0	8 3 | 8 4	6 6 | 5 2 | 0 9 | 0 7	11 2 | 8 7 | 2 1 | 1 6
14	*Scarelus*					26 9	39 3
15	several			7 0 | 1 1	13 4 | 3 5		
16	several	12 4	45 5	103 8	167 0	14 7 | 30 0 | 35 5 | 52 2 | 70 0	44 1 | 57 9 | 65 8 | 117 7 | 70 0
17	several	4 0 | 8 5	13 5 | 19 0	20 7 | 19 0	20 7 | 29 0		
18	*Macrobrachium*			1 9	4 0	0 0	0 5
19	*Macrobrachium*	3 2	9 6	15 9 | 1 8 | 4 6 | 8 2	25 8 | 6 9 | 12 8 | 15 1	16 5 | 0 0 | 6 2	28 3 | 0 6 | 15 8
20	several	11 5 | 29 2	34 6 | 66 5	11 5 | 1 2	34 6 | 5 7	2 4 | 14 7 | 3 6 | 7 3	8 5 | 35 6 | 14 8 | 21 9
21	several			14 5 | 17 3	14 5 | 17 3	14 5	14 5
22	*Corbicula*			0 8 | 0 9	0 8 | 2 0		
23	several	21 6	47 4	10 4 | 24 1	30 8 | 53 8		
24	*Sulcospira*			5 1 | 1 6	9 9 | 8 7		
25	*Barbodes*			1 3 | 0 3	14 6 | 5 7		
26	*Oryzias*			5 3	10 4	0 6	2 0
27	*Pangio*			17 8 | 1 4 | 15 3	35 2 | 3 8 | 22 7	0 5 | 20 8 | 11 5 | 0 7 | 2 1 | 4 1 | 6 8	2 0 | 33 2 | 19 3 | 1 9 | 6 1 | 9 9 | 16 7
28	*Limnonectes*						
29	*Rana*			6 6	13 0	4 0	8 5
30	*Rana*			7 0 | 1 6	13 3 | 2 9	2 8 | 1 4	5 0 | 2 5
31	several			10 5 | 7 9	20 2 | 17 0		
32	*Arachnothera*			3 6 | 0 2	6 8 | 0 6	0 2	0 6
33	*Cettia*	3 1	7 1	3 1	7 1		
34	*Ceyx*			0 1	0 3	0 1	0 3
35	*Copsychus*			0 3	0 9	0 2	0 7
36	*Dicaeum*					0 1	0 7
37	*Enicurus*			1 1	2 0	0 0	0 1
38	*Ficedula*	2 7	2 7				
39	*Lalage*						

		TMRCA IND		TMRCA A		TMRCA B	
		India/Sri Lanka	India/Sri Lanka	Indochina	Indochina	Malay peninsula	Malay peninsula
dataset	genus	min	max	min	max	min	max
40	*Nectarinia*					0 9	1 6
41	*Phylloscopus*						
42	*Pycnonotus*			1 9 | 2 3	5 6 | 6 6	0 1	2 0
43	*Rhipidura*			1 6 | 1 1	7 3 | 13 5	0 0 | 0 0	3 9 | 1 0
44	several						
45	several			16 6 | 1 8	21 3 | 2 4		
46	several						
47	several			4 8 | 4 9 | 9 6 | 3 0	8 3 | 8 8 | 14 5 | 5 7		
48	*Crocidura*	6 5	6 5	3 0 | 5 8	4 9 | 7 5	2 5 | 2 0	4 5 | 3 6
49	*Cuon*	0 4	1 5	0 3	1 1	0 2	0 2
50	*Cynopterus*	0 4	0 9	0 4	0 9	0 6	1 2
51	*Elaphas*	1 8 | 0 6	2 0 | 1 3	1 8 | 0 6	2 0 | 1 3	1 3	1 9
52	*Hipposideros*	2 9	2 9	4 8 | 1 2 | 2 1 | 1 9	9 3 | 3 1 | 4 2 | 3 3	0 6 | 0 3 | 0 2 | 0 4 | 0 1 | 0 6 | 0 2 | 1 2 | 2 3 | 2 0	1 7 | 1 0 | 0 5 | 0 8 | 0 5 | 1 2 | 0 6 | 3 1 | 4 7 | 3 6
53	*Macaca*	0 6	1 8	0 6 | 0 6 | 0 9	1 8 | 1 8 | 2 3	0 6	1 8
54	*Macaca*	0 9	1 7	0 9 | 4 0	1 7 | 6 0		
55	*Neofelis*			1 4	3 7	1 4	3 7
56	*Panthera*	0 2	1 7	0 2 | 0 2	1 5 | 1 5	0 2	1 5
57	*Paradoxurus*	2 8	5 8	2 5 | 1 1 | 3 2	2 5 | 2 9 | 3 2	3 2	3 2
58	several						
59	several			11 1	11 1		
60	*Sundascirus*					2 1 | 2 4	2 9 | 3 1
61	*Tupaia*			11 5	19 5		

	TMRCA C		TMRCA PAL		TMRCA D	
	Sumatra	Sumatra	Palawan	Palawan	Philippines	Philippines
dataset	min	max	min	max	min	max
01	13 6	37 4				
02					1 7 | 1 5 | 2 2 | 3 5 | 3 6	6 4 | 5 2 | 6 0 | 7 1 | 9 0
03	1 1 | 0 5	4 8 | 2 8			10 0 | 1 8 | 0 9	10 0 | 5 2 | 3 6
04			9 2 | 5 9 | 60 6	43 6 | 58 7 | 143 3	56 9 | 10 1 | 60 6	103 8 | 59 5 | 143 3
05	6 6| 4 7	6 6 | 11 3			0 9	5 6
06	6 0 | 1 5	10 0 | 2 0			9 0	14 0
07	0 3	1 0				
08	2 5 | 1 5	5 0 | 5 0			7 5 | 1 5	10 5 | 3 5
09	3 5	4 5	4 5 | 4 5	9 0 | 6 0	10 0 | 3 0	14 5 | 4 0
10			0 0	1 0	0 0 | 0 5 | 1 0 | 4 5 | 9 0 | 9 0 | 10 0	2 5 | 4 0 | 5 0 | 9 0 | 16 0 | 21 5
11	0 9 | 0 9	1 4 | 1 3			12 6	16 7
12					0 6 | 0 1	1 5 | 0 4
13	3 6 | 0 9 | 2 9	7 5 | 2 1 | 5 5				
14	26 9	39 3			22 1	35 1
15					2 4	5 2
16	27 8 | 4 1 | 63 9 | 25 3 | 43 7	68 1 | 30 5 | 100 1 | 56 5 | 43 7				
17						
18	0 0	0 4			0 8	2 0
19	3 6 | 2 6	13 4 | 8 1				
20	14 7	35 6			11 8	36 4
21	4 5 | 11 5	4 5 | 11 5			10 9	10 9
22	0 8	0 8				
23	10 4	30 8			31 6	59 1
24						
25	1 6	1 6				
26					0 6	1 9
27	0 5	2 0				
28					9 9 | 9 5	13 8 | 12 9
29	0 3 | 5 2 | 1 0	2 4 | 5 2 | 3 4				
30	3 4 | 1 4	6 4 | 2 5				
31					0 4	1 3
32			0 9	2 1	1 4 | 3 6	3 0 | 6 8
33	0 2	0 9				
34					2 5	4 0
35	0 0	0 2			0 7	2 1
36					3 1 | 4 0	6 2 | 9 3
37	0 0	0 1				
38			1 1	3 4	2 6 | 2 8	6 2 | 5 5
39	0 9	2 7			0 2	1 1

	TMRCA C		TMRCA PAL		TMRCA D	
	Sumatra	Sumatra	Palawan	Palawan	Philippines	Philippines
dataset	min	max	min	max	min	max
40	0 1	0 2			0 9	1 6
41			0 2	1 2	1 3	4 4
42	2 3 | 0 1	6 6 | 2 0			3 6 | 1 4	9 3 | 4 9
43	0 2	5 4			1 8 | 0 9	14 3 | 11 4
44						
45			5 7	8 0	9 0	11 6
46			0 2	0 9	0 8 | 0 2	1 9 | 0 9
47			3 8 | 3 0 | 6 0	7 9 | 5 9 | 11 4	4 8 | 5 1 | 3 0 | 4 9	8 3 | 10 4 | 5 7 | 7 2
48	2 0 | 4 2 | 1 7	4 6 | 5 8 | 2 8	4 5 | 1 7	6 7 | 2 6	3 3	3 3
49	0 0	0 3				
50			0 4 | 0 0	0 9 | 0 2	0 4	0 9
51	1 3	1 9				
52					2 1 | 0 6 | 0 6	4 8 | 1 5 | 1 5
53	0 0 | 0 3	0 4 | 1 0			0 2	0 8
54	2 9	4 7				
55						
56	0 2	1 5				
57	2 3 | 3 2	4 9 | 3 2	1 4	3 3	1 4	3 3
58					28 5 | 24 8 | 19 3 | 9 8 | 8 6	45 1 | 39 5 | 31 2 | 16 7 | 14 4
59						
60	2 8 | 4 7 | 2 4	3 7 | 6 3 | 3 1	4 9	7 0	4 1	5 9
61	3 5 | 5 0	9 5 | 12 5	9 0	16 0		

	TMRCA E		TMRCA F	
	Borneo	Borneo	Java	Java
dataset	min	max	min	max
01	82 | 29 | 41 | 36	264 | 179 | 171 | 173	17	134
02	98 | 11	180 | 53		
03	03	22		
04	569 | 152 | 849	1038 | 578 | 1491		
05	143 | 112	250 | 112	19	69
06	60 | 20	100 | 40	20	40
07	03	10	01	03
08	05	30	00 | 05	15 | 30
09	45 | 35	90 | 45	35 | 20	45 | 30
10	00 | 00 | 05	25 | 30 | 45	00 | 00 | 05 | 10 | 25 | 50	10 | 25 | 40 | 50 | 80 | 50
11	18 | 01 | 09	24 | 02 | 13	43	58
12				
13	66 | 36 | 63 | 12 | 12	127 | 75 | 63 | 26 | 26		
14	356	503	110	173
15	76 | 51	144 | 97		
16	930	1510	285 | 517	555 | 786
17	35	65		
18	01	08		
19	14	57	55 | 46	125 | 128
20	169 | 118 | 157 | 170	466 | 364 | 157 | 382	36 | 73	148 | 219
21	211	211	167	167
22			01	01
23			241	538
24	75 | 51	135 | 99	62	114
25			13	146
26			06	20
27	320	664		
28	118 | 95	169 | 129	118	169
29	21 | 17	66 | 46		
30	41 | 21	69 | 35	22	52
31	12	28		
32	97	145		
33	02	02	02	02
34	01	03	01	04
35	05	15		
36	04	13	01	07
37	10	17	08	16
38				
39	02	11		

	TMRCA E		TMRCA F	
	Borneo	Borneo	Java	Java
dataset	min	max	min	max
40				
41	0 6	3 0		
42	1 9 | 0 8	5 6 | 3 7		
43	0 9	11 4		
44				
45	3 5 | 0 7	5 0 | 1 3		
46	0 6	1 6		
47	8 4 | 5 1 | 3 0 | 5 9	13 6 | 10 4 | 5 9 | 11 4		
48	2 9	4 1	1 7	2 8
49			0 0	0 3
50	0 1 | 0 1 | 0 1	0 4 | 0 2 | 0 2		
51	1 3	1 9		
52	0 3 | 0 7 | 0 8 | 2 0 | 2 5 | 0 0	0 8 | 1 8 | 1 7 | 4 2 | 4 8 | 0 2		
53	0 2	0 8	0 0	0 4
54	0 7	1 4		
55	1 4	3 7		
56				
57	1 4	3 3	3 2	3 2
58				
59	11 3	11 3		
60	4 1 | 2 1 | 4 7 | 4 4	5 9 | 2 9 | 6 3 | 6 0		
61	3 5 | 12 5	9 5 | 21 5	5 0	13 5

	TMRCA G Wallacea Sulawesi	Wallacea Sulawesi
dataset	min	max
01	4 1	17 1
02	4 4 | 0 0 | 0 0	10 3 | 0 0 | 2 8
03	0 5 | 3 1 | 4 5 | 3 0	2 3 | 9 2 | 4 5 | 7 5
04		
05	7 8 | 3 6	7 8 | 11 0
06		
07		
08	21 5 | 6 0 | 13 5 | 1 0 | 2 5	28 0 | 13 5 | 22 0 | 3 0 | 5 0
09	8 5 | 10 0	14 0 | 16 0
10	0 0 | 0 5 | 0 5 | 1 0 | 1 0 | 1 5 | 1 5 | 2 0 | 2 0 | 3 0 | 3 5 | 4 0 | 5 0 | 5 5 | 6 0 | 7 0 | 7 0 |	8 0 | 3 0 | 4 5 | 5 0 | 5 5 | 4 5 | 5 0 | 5 5 | 6 5 | 8 5 | 10 5 | 11 0 | 5 0 | 11 0 | 12 0 | 7 0 | 13 5
11	13 7 | 8 9	17 0 | 11 4
12	0 1	0 1
13	3 9 | 6 5 | 1 7	8 3 | 10 5 | 3 6
14		
15		
16	32 6 | 43 9	37 4 | 77 4
17	2 0 | 11 0	9 0 | 22 0
18		
19	3 6 | 0 4 | 0 2	13 4 | 2 6 | 2 5
20	4 6 | 17 6 | 9 6 | 9 9	17 2 | 40 0 | 26 7 | 26 1
21		
22	1 2 | 0 8 | 0 1	1 2 | 0 8 | 0 1
23	20 0	42 0
24		
25		
26	3 6	6 5
27		
28	4 6 | 6 9 | 6 0	9 4 | 11 7 | 8 0
29		
30	5 3	9 2
31	1 2	2 8
32		
33		
34	3 8	7 3
35		
36	2 3	5 6
37		
38	2 6 | 3 6 | 0 0	6 2 | 3 6 | 0 3
39		

	TMRCA G Wallacea Sulawesi	
dataset	min	max
40	1 1	1 9
41		
42		
43		
44	6 6 | 3 3	13 8 | 8 4
45	5 9 | 1 4	7 8 | 2 1
46	2 4 | 1 7 | 0 7	5 2 | 3 9 | 1 8
47		
48	4 5 | 2 0	6 7 | 3 6
49		
50	0 4	1 0
51		
52	2 1 | 0 2 | 0 9 | 0 5 | 0 3	4 8 | 0 5 | 2 2 | 1 4 | 1 0
53		
54	2 9	4 7
55		
56		
57	3 2	3 2
58		
59	11 1	11 1
60		
61		

	TMRCA H E of Wallacea	E of Wallacea	
dataset	**min**	**max**	**remarks**
01	3 6 | 8 3	17 3 | 22 9	
02	3 6	9 0	not all TMRCA dates shown due to uncertain distribution
03	4 5	4 5	
04	121 6	121 6	
05	11 1	19 7	
06			
07			
08	3 0 | 6 0	9 0 | 11 0	
09			
10	0 0 | 1 5 | 2 0 | 4 5 | 6 0 | 7 0 | 8 0 | 10 0 | 13 5 | 21 0	8 0 | 6 0 | 6 5 | 10 5 | 6 0 | 13 5 | 8 0 | 20 5 | 22 0 | 21 0	
11	18 8	23 6	
12	0 3 | 0 0	0 8 | 0 3	
13	5 0	8 9	
14			
15	1 1 | 2 3 | 3 5	3 5 | 4 9 | 6 4	
16	1 6	12 5	
17	2 0 | 19 0 | 5 5 | 17 5 | 28 5	9 0 | 29 0 | 7 5 | 29 5 | 44 0	
18	1 9	4 0	
19	1 8	6 9	not all TMRCA dates shown due to low branch supports for terminal splits
20	6 3 | 12 8	20 | 30 9	
21			
22	0 8	0 8	
23	20 0	42 0	
24			
25			
26			
27			
28			
29			
30			
31	8 3	16 2	
32			
33			
34	1 7	3 1	
35			
36	2 3	5 6	
37			
38			not all TMRCA dates shown due to uncertain distribution
39	2 2	5 6	

dataset	TMRCA H E of Wallacea min	E of Wallacea max	remarks
40	0 1	0 2	
41	0 6	3 0	Phylloscopus trivirgatus only
42			
43	2 3	20 2	
44	6 6 | 2 0 | 0 3	13 8 | 5 5 | 1 2	
45	1 9	2 4	not all TMRCA dates shown due to low branch supports for terminal splits
46	3 9	9 4	
47			
48			data shown for third tree (fossil + rates used p 177)
49			
50			not all TMRCA dates shown due to low branch supports for terminal splits
51			
52			not all TMRCA dates shown due to low branch supports for terminal splits
53			
54			
55			
56			
57			
58			not all TMRCA dates shown due to uncertain distribution
59			not all TMRCA dates shown due to uncertain distribution
60			
61			

Table S2. Brief summary of BEAST results showing divergence times for root and oldest MRCA between complex (selected by jModeltest) and less complex models (*a posteriori* selected due to low ESS values; see text for details).

Dataset	nTaxa ingroup	Model used (see comments in Table 2)	Root height mean (model used)	Oldest MRCA 95% HPD	Root height Mean (previously used model)	Oldest MRCA 95% HPD (previously used model)
04	30	HKY+G	167.94	96.96 | 178.59	166.36 (GTR+G)	96.99 | 178.02 (GTR+G)
24	129	GTR+G	13.17	8.40 | 15.81	13.60 (HKY+G)	8.77 | 15.77 (HKY+G)
30	54	HKY+I+G	11.39	6.99 | 13.27	11.80 (GTR+G)	7.26 | 14.09 (GTR+G)
32	46	HKY+I+G | HKY+I+G	15.71	9.73 | 14.50	16.28 (GTR+G | GTR+G)	9.65 | 15.59 (GTR+G | GTR+G)
35	33	HKY+I+G | HKY+I	1.75	0.65 | 2.07	1.98 (HKY+G | HKY+G)	0.74 | 2.29 (HKY+G | HKY+G)
42	46	HKY+I+G	8.39	3.97 | 11.10	8.59 (HKY+G)	3.87 | 11.62 (HKY+G)
43	50	HKY+I+G	9.56	2.25 | 20.18	6.52 (HKY+G)	1.87 | 12.39 (HKY+G)
50	140	HKY+I+G	1.23	0.68 | 1.30	1.29 (GTR+G)	0.74 | 1.43 (GTR+G)
56	25	HKY+I+G	1.58	0.33 | 1.47	1.16 (HKY+G)	0.27 | 1.24 (HKY+G)

Table S3. Simplified statistics for 'Out-of-' colonization routes among all taxa groups and *a priori* defined timeframes, with detailed information on Borneo.

Area overview (among all taxa roups and timeframes)	**Total**	**% of 49 datasets[a]**	**Additive %[b]**		
N(datasets) out of A (Indochina)	36	73.5	24.5		
N(datasets) out of B (Thai-Malay Peninsula)	12	24.5	8.2		
N(datasets) out of C (Sumatra)	14	28.6	9.5		
N(datasets) out of D (Philippines)	12	24.5	8.2		
N(datasets) out of PAL (Palawan)	2	4.1	1.4		
N(datasets) out of E (Borneo)	32	65.3	21.8		
N(datasets) out of F (Java)	7	14.3	4.8		
N(datasets) out of G (Wallacea excl. Sulawesi)	5	10.2	3.4		
N(datasets) out of SUL (Sulawesi)	9	18.4	6.1		
N(datasets) out of H (East of Wallace Line)	12	24.5	8.2		
N(datasets) out of IND (India+Sri Lanka)	4	8.2	2.7		
N(datasets) out of JAP (Japan)	0	0	0		
Borneo in detail (OoB = Out-of-Borneo)		**Total**	**pre-Miocene**	**Miocene**	**Plio-Pleistocene**
	plants	4			
N(datasets) OoB | no. of OoB colonizations		3	1|3	3|9	2|2
dataset | no. of OoB colonizations			04|3	01|2 02|1 05|6	02|1 05|1
N(datasets) studied	insects/spiders	6			
N(datasets) OoB | no. of OoB colonizations		6	2|3	2|7	3|4
dataset | no. of OoB colonizations			14|1 16|2	13|5 15|2	09|1 11|2 13|1
N(datasets) studied	crustaceans	4			
N(datasets) OoB | no. of OoB colonizations		2	1|4	2|5	
dataset | no. of OoB colonizations			21|4	20|2 21|3	
N(datasets) studied	molluscs	3			
N(datasets) OoB | no. of OoB colonizations		1	0	1|1	0
dataset | no. of OoB colonizations				24|1	
N(datasets) studied	fishes	3			
N(datasets) OoB | no. of OoB colonizations		1	1|2	1|2	0
dataset | no. of OoB colonizations			27|2	27|2	
N(datasets) studied	herpetofauna	4			
N(datasets) OoB | no. of OoB colonizations		4	0	2|4	2|4
dataset | no. of OoB colonizations				28|2 29|2	30|3 31|1
N(datasets) studied	birds	16			
N(datasets) OoB | no. of OoB colonizations		8	0	3|6	6|12
dataset | no. of OoB colonizations				32|2 45|1 47|3	32|2 35|3 36|1 37|3 42|2 43|1
N(datasets) studied	mammals	9			
N(datasets) OoB | no. of OoB colonizations		6	0	1|3	5|14
dataset | no. of OoB colonizations				61|3	48|3 52|3 55|2 57|2 60|4

[a] In only 49 out of 61 datasets, ancestral area reconstructions, and thus colonization route reconstructions were performed in the present study or in the source publication (see comments in Table 1).
[b] Additive percentage of total 147 colonization routes observed among all taxa groups and timeframes (redundant 'Out-of' colonizations in a single dataset were counted once only).

Table S4 (next pages). Mann-Whitney *U* test results for total number of lineages, *in situ* diversifications, and number of emigrations and immigrations identified among all areas across all taxonomic groups (see text for details). Area codes denote biogeographic areas defined *a priori*: A=Indochina, B=Thai-Malay Peninsula, C=Sumatra, D=Philippines, PAL=Palawan, E=Borneo, F=Java, G=Wallacea excluding Sulawesi, SUL=Sulawesi, H=East of Wallacea. Significance levels: * $p<0.05$, ** $p<0.01$, and *** $p<0.001$. Yellow-marked fields denote significant differences found after Bonferroni correction, i.e. $p=0.05$/number of tests (see text for details).

Total number of lineages

	A	B	C	D	PAL	E	F	G	SUL	H
Plants	35	15	19	43	9	65	6	5	38	25
Insects/spiders	77	47	58	41	7	56	32	47	30	123
Crustaceans	91	25	8	19	0	34	6	1	6	51
Molluscs	44	0	6	2	0	9	18	0	31	3
Fishes	44	22	5	1	0	21	4	0	5	0
Herpetofauna	28	13	9	31	0	28	4	1	44	3
Birds	61	25	15	123	13	100	10	17	13	89
Mammals	144	86	48	119	17	49	13	3	32	1

Total number of lineages excl. pop-based datasets

	A	B	C	D	PAL	E	F	G	SUL	H
Plants	35	15	19	43	9	65	6	5	38	25
Insects/spiders	39	38	43	27	7	53	30	44	30	118
Crustaceans	66	19	7	6	0	31	6	1	6	7
Molluscs	44	0	6	2	0	9	18	0	31	3
Fishes	31	22	2	1	0	21	1	0	5	0
Herpetofauna	28	13	9	31	0	28	4	1	44	3
Birds	54	10	9	52	12	58	7	17	9	70
Mammals	79	62	27	112	13	44	8	1	32	1

Mann-Whitney U test: Ntotal number of lineages

	A	B	C	D	PAL	E	F	G	SUL	H
A										
B	*									
C	*									
D										
PAL	**	*		*						
E			*		**					
F	**					**				
G	**		*	*		**				
SUL	*				*			*		
H										

Mann-Whitney U test: Total number of lineages excl. pop-based datasets

	A	B	C	D	PAL	E	F	G	SUL	H
A										
B	*									
C	**									
D										
PAL	**	*		*						
E			*		**					
F	**					**				
G	**			*		**				
SUL	*				*			*		
H										

Total number of in situ diversifications

	A	B	C	D	PAL	E	F	G	SUL	H
Plants	29	2	7	31	6	58	2	0	29	27
Insects/spiders	56	28	35	22	3	34	17	23	13	92
Crustaceans	79	14	1	17	0	22	1	0	0	47
Molluscs	37	0	5	0	0	6	14	0	27	1
Fishes	40	11	3	0	0	14	1	0	4	0
Herpetofauna	25	9	6	26	0	22	2	0	39	1
Birds	30	9	6	6	2	41	2	0	2	1
Mammals	107	56	28	41	7	22	5	1	23	0

Total number of in situ diversifications excl. pop-based datasets

	A	B	C	D	PAL	E	F	G	SUL	H
Plants	29	2	7	31	6	58	2	0	29	27
Insects/spiders	22	23	23	12	3	32	16	21	13	92
Crustaceans	59	12	1	5	0	20	1	0	0	47
Molluscs	37	0	5	0	0	6	14	0	27	1
Fishes	29	11	1	0	0	14	0	0	4	0
Herpetofauna	25	9	6	26	0	22	2	0	39	1
Birds	33	4	1	37	5	40	0	5	2	1
Mammals	56	39	12	36	7	19	3	0	23	0

Mann-Whitney U test: Ntotal in situ diversifications

	A	B	C	D	PAL	E	F	G	SUL	H
A										
B	**									
C	**									
D	*									
PAL	**	*	*	*						
E			*		**					
F	**					**				
G	**	*	**	*		**	*			
SUL	**				*			*		
H	*							*		

Mann-Whitney U test: Ntotal in situ diversifications excl. pop-based datasets

	A	B	C	D	PAL	E	F	G	SUL	H
A										
B	**									
C	**									
D										
PAL	**	*								
E			*		**					
F	**					**				
G	**	*	*	*		**				
SUL	*				*			*		
H										

Proportion: in situ diversifications/lineages (excl. pop-based datasets)

	A	B	C	D	PAL	E	F	G	SUL	H
Plants	0.83	0.13	0.37	0.72	0.67	0.89	0.33		0.76	1.08
Insects/spiders	0.56	0.61	0.53	0.44	0.43	0.60	0.53	0.48	0.43	0.78
Crustaceans	0.89	0.63	0.14	0.83		0.65	0.17			6.71
Molluscs	0.84		0.83			0.67	0.78		0.87	0.33
Fishes	0.94	0.50	0.50			0.67			0.80	
Herpetofauna	0.89	0.69	0.67	0.84		0.79	0.50		0.89	0.33
Birds	0.61	0.40	0.11	0.71	0.42	0.69		0.29	0.22	0.01
Mammals	0.71	0.63	0.44	0.32	0.54	0.43	0.38		0.72	

Mann-Whitney U test: in situ diversifications/lineages

	A	B	C	D	PAL	E	F	G	SUL	H
A										
B	*									
C	**									
D										
PAL	*									
E										
F	**					*				
G	*									
SUL										
H										

Total number of emigrations

	A	B	C	D	PAL	E	F	G	SUL	H
Plants	8	5	2	2	2	16	0	5	4	3
Insects/spiders	9	9	6	7	0	14	4	7	5	6
Crustaceans	11	1	2	1	0	9	1	0	0	1
Molluscs	8	0	0	0	0	1	1	0	1	1
Fishes	10	1	0	0	0	4	0	0	0	0
Herpetofauna	4	2	2	1	0	8	1	0	2	1
Birds	20	2	4	16	0	19	1	2	2	8
Mammals	25	5	7	2	1	18	0	0	1	0

Total number of immigrations

	A	B	C	D	PAL	E	F	G	SUL	H
Plants	1	4	4	8	1	2	3	0	0	5
Insects/spiders	7	5	6	5	1	6	4	2	5	6
Crustaceans	3	4	2	3	0	2	2	0	4	2
Molluscs	2	0	1	0	0	0	1	0	1	1
Fishes	2	3	2	1	0	0	1	0	1	0
Herpetofauna	1	3	3	3	0	2	3	1	3	1
Birds	6	7	6	9	5	12	4	4	8	7
Mammals	9	11	8	7	6	9	5	0	5	0

Mann-Whitney U test: Ntotal emigrations

	A	B	C	D	PAL	E	F	G	SUL	H
A										
B	**									
C	**									
D	**									
PAL	**	*	*	*						
E		*	*	*	**					
F	**					**				
G	**					**				
SUL	**				*	**				
H	**				*	**				

Mann-Whitney U test: Ntotal immigrations

	A	B	C	D	PAL	E	F	G	SUL	H
A										
B										
C										
D										
PAL	*		*							
E										
F										
G	*	*	**	*			*			
SUL								*		
H										

Mann-Whitney U test: Total number of emigrations vs. immigrations

A	**
E	*
F	*

Total number of emigrations (between time frames)

	pre	Mio	PP
A	2	30	63
B	7	7	11
C	3	6	14
D	0	12	17
PAL	2	1	0
E	12	41	36
F	2	3	3
G	0	5	9
SUL	1	10	4
H	1	7	12

Total number of immigrations (between time frames)

	pre	Mio	PP
A	2	13	16
B	5	10	22
C	3	10	19
D	5	8	23
PAL	1	5	7
E	2	9	22
F	1	9	13
G	0	3	4
SUL	3	11	13
H	4	8	10

Mann-Whitney U test: Ntotal emigrations

Pre-Miocene vs. Miocene	
Pre-Miocene vs. Plio-Pleistocene	**
Miocene vs. Plio-Pleistocene	

Mann-Whitney U test: Ntotal immigrations

Pre-Miocene vs. Miocene	**
Pre-Miocene vs. Plio-Pleistocene	***
Miocene vs. Plio-Pleistocene	*

Proportion: Ntotal emigrations/Ntotal lineages excl. pop-based

	A	B	C	D	PAL	E	F	G	SUL	H
Plants	0.23	0.33	0.11	0.05	0.22	0.25		1.00	0.11	0.12
Insects/spiders	0.23	0.24	0.14	0.26		0.26	0.13	0.16	0.17	0.05
Crustaceans	0.17	0.05	0.29	0.17		0.29	0.17			0.14
Molluscs	0.18					0.11	0.06		0.03	0.33
Fishes	0.32	0.05				0.19				
Herpetofauna	0.14	0.15	0.22	0.03		0.29	0.25		0.05	0.33
Birds	0.37	0.20	0.44	0.31		0.33	0.14	0.12	0.22	0.11
Mammals	0.32	0.08	0.26	0.02	0.08	0.41			0.03	

Mann-Whitney U test: Ntotal emigrations/Ntotal lineages

	A	B	C	D	PAL	E	F	G	SUL	H
A										
B										
C										
D										
PAL										
E										
F						*				
G										
SUL	**		*			**				
H										

Non-used datasets

Alfaro M.E., Karns D.R., Voris H.K., Abernathy E., Sellins S.L. 2004. Phylogeny of *Cerberus* (Serpentes: Homalopsinae) and phylogeography of *Cerberus rynchops*: diversification of a coastal marine snake in Southeast Asia. J. Biogeogr. 31:1277–1292.

- see dataset 31

Azuma N., Ogata K., Kikuchi T., Higashi S. 2006. Phylogeography of Asian weaver ants, *Oecophylla smaragdina*. Ecol. Res. 21:126–136.

- sequences not available
- authors using a rate of 0.165%/Ma for 1st and 2nd position → max. age of Asian lineages: 13.3 Mya

Balke M., Ribera I. 2004. Jumping across Wallace's Line: *Allodessus* Guignot and *Limbodessus* Guignot revisited (Coleoptera: Dytiscidae, Bidessini) based on molecular-phylogenetic and morphological data. Aust. J. Entomol. 43:114–128.

- see water beetle datasets in Stelbrink et al. (2012)

Bänfer G., Moog U., Fiala B., Mohamed M.A., Weising K., Blattner F.R. 2006. A chloroplast genealogy of myrmecophytic *Macaranga* species (Euphorbiaceae) in Southeast Asia reveals hybridization, vicariance and long-distance dispersals. Mol. Ecol. 15:4409–4424.

- outgroup missing
- authors: max. age of Asian lineages: 18,000 years

Blackburn D.C., Bickford D.P., Diesmos A.C., Iskandar D.T., Brown R.M. 2010. An ancient origin for the enigmatic flat-headed frogs (Bombinatoridae: *Barbourula*) from the islands of Southeast Asia. PLoS ONE. 5:e12090.

- only Borneo and Palawan
- authors: age of Asian lineages: Miocene-Paleocene

Buerki S., Forest F., Alvarez N., Nylander J.A.A., Arrigo N., Sammartín I. 2011. An evaluation of new parsimony-based versus parametric inference methods in biogeography: a case study using the globally distributed plant family Sapindaceae. J. Biogeogr. 38:531–550.

- high-level phylogeny (several genera and families)
- authors: first diversifications in the Cretaceous

Butlin R.K., Walton C., Monk K.A., Bridle J.R., Hall R., Holloway J.D. 1998. Biogeography of Sulawesi grasshoppers, genus *Chitaura*, using DNA sequence data. In: Hall R., Holloway D. Biogeography and Geological Evolution of SE Asia. Leiden: Backhuys Publishers. p. 355–359.

- see grasshopper dataset in Stelbrink et al. (2012)
- only Sulawesi taxa

Campbell P., Schneider C.J., Adnan A.M., Zubaid A., Kunz T.H. 2006. Comparative population structure of *Cynopterus* fruit bats in peninsular Malaysia and southern Thailand. Mol. Ecol. 15:29–47.

- see dataset 50

Cannon C.H., Manos P.S. 2003. Phylogeography of the Southeast Asian stone oaks (*Lithocarpus*). J. Biogeogr. 30:211–226.

- only Thailand and Borneo
- authors: age of Asian lineages: Miocene

Che J., Zhou W.-W., Hu J.-S., Yan F., Papenfuss T.J., Wake D.B., Zhang Y.-P. 2010. Spiny frogs (Paini) illuminate the history of the Himalayan region and Southeast Asia. Proc. Natl. Acad. Sci. USA. 107:13765–13770.
- only mainland SE Asia
- authors: age of Asian lineages: Miocene

Cox M.P., Karafet T.M., Lansing J.S., Sudoyo H., Hammer M.F. 2010. Autosomal and X-linked single nucleotide polymorphisms reveal a steep Asian-Melanesian ancestry cline in eastern Indonesia and a sex bias in admixture rates. Proc. R. Soc. London B 277:1589–1596.
- human dataset
- authors: age of Asian lineages: Pleistocene

Das A., Mohanty S., Stephan W. 2004. Inferring the population structure and demography of *Drosophila ananassae* from multilocus data. Genetics. 168:1975–1985.
- outgroup missing
- authors: age of Sundaland lineages: 18,000 years

de Bruyn M., Nugroho E., Hossain M.M., Wilson J.C., Mather P.B. 2005. Phylogeographic evidence for the existence of an ancient biogeographic barrier: the Isthmus of Kra Seaway. Heredity 94:370–378.
- see dataset 19

de Bruyn M., Wilson J.A., Mather P.B. 2004. Huxley's Line demarcates extensive genetic divergence between eastern and western forms of the giant freshwater prawn, *Macrobrachium rosenbergii*. Mol. Phylogenet. Evol. 30:251–257.
- see dataset 19

Esselstyn J.A., Oliveros C.H. 2010. Colonization of the Philippines from Taiwan: a multi-locus test of the biogeographic and phylogenetic relationships of isolated populations of shrews. J. Biogeogr. 37:1504–1514.
- see dataset 48

Esselstyn J.A., Oliveros C.H., Moyle R.G., Peterson A.T., McGuire J.A., Brown R.M. 2010. Integrating phylogenetic and taxonomic evidence illuminates complex biogeographic patterns along Huxley's modification of Wallace's Line. J. Biogeogr. 37:2054–2066.
- different groups analyzed: birds, mammals, amphibians, lizards, snakes, and turtles
- review showing sister group relationship to Palawan taxa

Esselstyn J.A., Timm R.M., Brown R.M. 2009. Do geological or climatic processes drive speciation in dynamic archipelagos? The tempo and mode of diversification in Southeast Asian shrews. Evolution 63:2595–2610.
- only Philippine taxa
- authors: age of Asian lineages: Pliocene-Miocene

Fleischer R.C., Perry E.A., Muralidharan K., Stevens E.E., Wemmer C.M. 2001. Phylogeography of the asian elephant (*Elephas maximus*) based on mitochondrial DNA. Evolution 55:1882–1892.
- see dataset 51

Francis C.M., Borisenko A. V, Ivanova N. V, Eger J.L., Lim B.K., Guillén-Servent A., Kruskop S. V, Mackie I., Hebert P.D.N. 2010. The role of DNA barcodes in understanding and conservation of mammal diversity in Southeast Asia. PLoS ONE. 5:e12575.
- many taxa (1896 sequences), and distribution information not available (only in GenBank)

Gaubert P., Denys G., Oberdorff T. 2009. Genus-level supertree of Cyprinidae (Actinopterygii: Cypriniformes), partitioned qualitative clade support and test of macro-evolutionary scenarios. Biol. Rev. 84:653–689.
- many taxa, but distribution information not sufficient

Gongora J., Rawlence N.J., Mobegi V., Jianlin H., Alcalde J., Matus J.T., Hanotte O., Moran C., Austin J.J., Ulm S., Anderson A.J., Larson G., Cooper A. 2008. Indo-European and Asian origins for Chilean and Pacific chickens revealed by mtDNA. Proc. Natl. Acad. Sci. USA. 105:10308–10313.
- domestic chickens without interesting pattern in SE Asia

Gorog A.J., Sinaga M.H., Engstrom M.D. 2004. Vicariance or dispersal? Historical biogeography of three Sunda shelf murine rodents (*Maxomys surifer*, *Leopoldamys sabanus* and *Maxomys whiteheadi*). Biol. J. Linn. Soc. 81:91–109.
- sequences not available
- authors: max. age of Asian lineages: 6 Mya

Guicking D., Fiala B., Blattner F.R., Slik F., Mohamed M., Weising K. 2011. Comparative chloroplast DNA phylogeography of two tropical pioneer trees, *Macaranga gigantea* and *Macaranga pearsonii* (Euphorbiaceae). Tree Genet. Genomes. 7:573–585.
- outgroup missing
- only Thai-Malay Peninsula and Borneo

Hahn W.J., Sytsma K.J. 1999. Molecular systematics and biogeography of the Southeast Asian genus *Caryota* (Palmae). Syst. Bot. 24:558–580.
- origin uncertain due to cultivated specimens
- authors: *C. mitis* on Sulawesi 10-25 Mya (based on geology)

Hu J., Zhang J.L., Nardi F., Zhang R.J. 2008. Population genetic structure of the melon fly, *Bactrocera cucurbitae* (Diptera: Tephritidae), from China and Southeast Asia. Genetica. 134:319–324.
- few sample sites:
- authors: authors: age of Asian lineages: Pleistocene?

Javadi F., Tun Y.T., Kawase M., Guan K., Yamaguchi H. 2011. Molecular phylogeny of the subgenus *Ceratotropis* (genus *Vigna*, Leguminosae) reveals three eco-geographical groups and Late Pliocene-Pleistocene diversification: evidence from four plastid DNA region sequences. Ann. Bot. 108:367–380.
- mostly mainland SE Asia
- authors: age of Asian lineages: Pleistocene-Pliocene

Jønsson K.A., Fabre P.H., Ricklefs R.E., Fjeldså J. 2011. Major global radiation of corvoid birds originated in the proto-Papuan archipelago. Proc. Natl. Acad. Sci. USA. 108:2328–2333.
- high-level phylogeny with origin in New Guinea
- authors: age of Asian lineages: Miocene-Oligocene

Karns D.R., O'Bannon A., Voris H.K., Weigt L.A. 2000. Biogeographical implications of mitochondrial DNA variation in the Bockadam snake (*Cerberus rynchops*, Serpentes: Homalopsinae) in Southeast Asia. J. Biogeogr. 27:391–402.
- see dataset 31

Köhler F., Glaubrecht M. 2007. Out of Asia and into India: on the molecular phylogeny and biogeography of the endemic freshwater gastropod *Paracrostoma* Cossmann, 1900 (Caenogastropoda: Pachychilidae). Biol. J. Linn. Soc. 91:627–651.
- see dataset 23

Kumazawa Y., Nishida M. 2000. Molecular phylogeny of osteoglossoids: a new model for Gondwanian origin and plate tectonic transportation of the Asian arowana. Mol. Biol. Evol. 17:1869–1878.
- origin of specimens uncertain
- authors: colonization of Asia: 20-170 Mya (depending on scenario)

Kurabayashi A., Matsui M., Belabut D.M., Yong H.-S., Ahmad N., Sudin A., Kuramoto M., Hamidy A., Sumida M. 2011. From Antarctica or Asia? New colonization scenario for Australian-New Guinean narrow mouth toads suggested from the findings on a mysterious genus *Gastrophrynoides*. BMC Evol. Biol. 11:175.
- origin of specimens uncertain
- authors: colonization of island SE Asia: 25 Mya

Larson G., Cucchi T., Fujita M., Matisoo-Smith E., Robins J., Anderson A., Rolett B., Spriggs M., Dolman G., Kim T.H., Thuy N.T.D., Randi E., Doherty M., Due R.A., Bollt R., Djubiantono T., Griffin B., Intoh M., Keane E., Kirch P., Li K.T., Morwood M., Pedriña L.M., Piper P.J., Rabett R.J., Shooter P., Van den Bergh G., West E., Wickler S., Yuan J., Cooper A., Dobney K. 2007. Phylogeny and ancient DNA of *Sus* provides insights into neolithic expansion in Island Southeast Asia and Oceania. Proc. Natl. Acad. Sci. USA. 104:4834–4839.
- see pigs dataset in Stelbrink et al. (2012)

Larson G., Dobney K., Albarella U., Fang M., Matisoo-Smith E., Robins J., Lowden S., Finlayson H., Brand T., Willerslev E., Rowley-Conwy P., Andersson L., Cooper A. 2005. Worldwide phylogeography of wild boar reveals multiple centers of pig domestication. Science 307:1618–1621.
- see pigs dataset in Stelbrink et al. (2012)

Laurent S.J.Y., Werzner A., Excoffier L., Stephan W. 2011. Approximate Bayesian analysis of *Drosophila melanogaster* polymorphism data reveals a recent colonization of Southeast Asia. Mol. Biol. Evol. 28:2041–2051.
- outgroup missing

Les D.H., Crawford D.J., Kimball R.T., Moody M.L., Landolt E. 2003. Biogeography of discontinuously distributed hydrophytes: a molecular appraisal of intercontinental disjunctions. Int. J. Plant Sci. 164:917–932.
- patchy wordlwide distribution; Asia not well represented
- authors: age of Asian lineages: 10.7/20.5 Mya

Liao P., Chiang Y., Huang S., Wang J. 2009. Gene flow of *Ceriops tagal* (Rhizophoraceae) across the Kra Isthmus in the Thai Malay Peninsula. Bot. Stud. 50:193–204.
- outgroup missing
- only Kra region

Liao P.-C., Havanond S., Huang S. 2007. Phylogeography of *Ceriops tagal* (Rhizophoraceae) in Southeast Asia: the land barrier of the Malay Peninsula has caused population differentiation between the Indian Ocean and South China Sea. Conserv. Genet. 8:89–98.
- outgroup missing
- authors: dispersal to East Asia: 16 Mya

Lim H.C., Zou F., Taylor S.S., Marks B.D., Moyle R.G., Voelker G., Sheldon F.H. 2010. Phylogeny of magpie-robins and shamas (Aves: Turdidae: *Copsychus* and *Trichixos*): implications for island biogeography in Southeast Asia. J. Biogeogr. 37:1894–1906.
- see dataset 35

Liu Y.P., Wu G.S., Yao Y.G., Miao Y.W., Luikart G., Baig M., Beja-Pereira A., Ding Z.L., Palanichamy M.G., Zhang Y.P. 2006. Multiple maternal origins of chickens: out of the Asian jungles. Mol. Phylogenet. Evol. 38:12–19.
- domestic chickens without interesting pattern in SE Asia

Lohman D.J., Peggie D., Pierce N.E., Meier R. 2008. Phylogeography and genetic diversity of a widespread Old World butterfly, *Lampides boeticus* (Lepidoptera: Lycaenidae). BMC Evol. Biol. 8:301.
- population-based dataset; original topology in unresolved
- authors: age of Old World lineages: 1.5 Mya

Loo A.H.B., Dransfield J., Chase M.W., Baker W.J. 2006. Low-copy nuclear DNA, phylogeny and the evolution of dichogamy in the betel nut palms and their relatives (Arecinae; Arecaceae). Mol. Phylogenet. Evol. 39:598–618.
- origin of specimens uncertain
- no rate availabe for nuDNA
- authors: age of New Guinean taxa: 15 Mya

Lovejoy N.R., Mullen S.P., Sword G.A., Chapman R.F., Harrison R.G. 2006. Ancient trans-Atlantic flight explains locust biogeography: molecular phylogenetics of *Schistocerca*. Proc. R. Soc. London B 273:767–774.
- only two sister taxa
- authors: post-dating Gondwanan break-up

Lucchini V., Meijaard E., Diong C.H., Groves C.P., Randi E. 2005. New phylogenetic perspectives among species of South-east Asian wild pig (*Sus* sp.) based on mtDNA sequences and morphometric data. J. Zool. 266:25–35.
- see pig dataset in Stelbrink et al. (2012)

Macaulay V., Hill C., Achilli A., Rengo C., Clarke D., Meehan W., Blackburn J., Semino O., Scozzari R., Cruciani F., Taha A., Shaari N.K., Raja J.M., Ismail P., Zainuddin Z., Goodwin W., Bulbeck D., Bandelt H.-J., Oppenheimer S., Torroni A., Richards M. 2005. Single, rapid coastal settlement of Asia revealed by analysis of complete mitochondrial genomes. Science 308:1034–1036.
- human dataset
- authors: age of Asian lineages: Pleistocene

Macey J.R., Schulte J.A., Larson A., Ananjeva N.B., Wang Y., Pethiyagoda R., Rastegar-Pouyani N., Papenfuss T.J. 2000. Evaluating trans-Tethys migration: an example using acrodont lizard phylogenetics. Syst. Biol. 49:233–256.
- insufficient distribution
- authors: age of Asian lineages: 160 Mya

Matsui M., Hamidy A., Murphy R.W., Khonsue W., Yambun P., Shimada T., Ahmad N., Belabut D.M., Jiang J.-P. 2010a. Phylogenetic relationships of megophryid frogs of the genus *Leptobrachium* (Amphibia, Anura) as revealed by mtDNA gene sequences. Mol. Phylogenet. Evol. 56:259–272.
- excluded due to reviewer's comment: calibration questionable (overestimation)
- authors: max. age of Asian lineages: Eocene

Matsui M., Tominaga A., Liu W., Khonsue W., Grismer L.L., Diesmos A.C., Das I., Sudin A., Yambun P., Yong H., Sukumaran J., Brown R.M. 2010b. Phylogenetic relationships of *Ansonia* from Southeast Asia inferred from mitochondrial DNA sequences: systematic and biogeographic implications (Anura: Bufonidae). Mol. Phylogenet. Evol. 54:561–570.
- excluded due to reviewer's comment: calibration questionable (overestimation)
- authors: max. age of Asian lineages: 74.9 Mya

Murphy W.J., Collier G.E. 1997. A molecular phylogeny for aplocheiloid fishes (Atherinomorpha, Cyprinodontiformes): the role of vicariance and the origins of annualism. Mol. Biol. Evol. 14:790–799.
- only two taxa

Nater A., Nietlisbach P., Arora N., Schaik C.P. van, Noordwijk M.A. van, Willems E.P., Singleton I., Wich S.A., Goossens B., Warren K.S., Verschoor E.J., Perwitasari-Farajallah D., Pamungkas J., Krützen M. 2011. Sex-biased dispersal and volcanic activities shaped phylogeographic patterns of extant Orangutans (genus: *Pongo*). Mol. Biol. Evol. 28:2275–2288.
- only Sumatra and Borneo

O'Loughlin S.M., Okabayashi T., Honda M., Kitazoe Y., Kishino H., Somboon P., Sochantha T., Nambanya S., Saikia P.K., Dev V., Walton C. 2008. Complex population history of two *Anopheles dirus* mosquito species in Southeast Asia suggests the influence of Pleistocene climate change rather than human-mediated effects. J. Evol. Biol. 21:1555–1569.
- only mainland SE Asia
- authors: age of Asian lineages: Pleistocene

Päckert M., Martens J., Sun Y.-H., Severinghaus L.L., Nazarenko A.A., Ting J., Töpfer T., Tietze D.T. 2012. Horizontal and elevational phylogeographic patterns of Himalayan and Southeast Asian forest passerines (Aves: Passeriformes). J. Biogeogr. 39:556–573.
- few samples from SE Asia

Pouyard L., Teugels G.G., Gustiano R., Legendre M. 2000. Contribution to the phylogeny of pangasiid catfishes based on allozymes and mitochondrial DNA. J. Fish Biol. 56:1509–1538.
- sequences not available
- author: max. age of Asian lineages: 19 Mya

Quek S.-P., Davies S.J., Ashton P.S., Itino T., Pierce N.E. 2007. The geography of diversification in mutualistic ants: a gene's-eye view into the Neogene history of Sundaland rain forests. Mol. Ecol. 16:2045–2062.
- host relationship
- authors: max. age of Asian lineages: Miocene

Quek S.-P., Davies S.J., Itino T., Pierce N.E. 2004. Codiversification in an ant-plant mutualism: stem texture and the evolution of host use in *Crematogaster* (Formicidae: Myrmicinae) inhabitants of *Macaranga* (Euphorbiaceae). Evolution 58:554–570.
- host relationship
- authors: max. age of Asian lineages: 6-12 Mya

Roberts T.E. 2006. History, ocean channels, and distance determine phylogeographic patterns in three widespread Philippine fruit bats (Pteropodidae). Mol. Ecol. 15:2183–2199.
- only Philippines

Roelants K., Gower D.J., Wilkinson M., Loader S.P., Biju S.D., Guillaume K., Moriau L., Bossuyt F. 2007. Global patterns of diversification in the history of modern amphibians. Proc. Acad. Nat. Sci. USA 104:887–892.
- complex high-level phylogeny among amphibians

Roelants K., Jiang J., Bossuyt F. 2004. Endemic ranid (Amphibia: Anura) genera in southern mountain ranges of the Indian subcontinent represent ancient frog lineages: evidence from molecular data. Mol. Phylogenet. Evol. 31:730–740.
- excluded due to reviewer's comment: several taxa missing in dataset
- authors: age of Asian lineages: 20-60 Mya

Rosenblum L.L., Supriatna J., Melnick D.J. 1997. Phylogeographic analysis of pigtail macaque populations (*Macaca nemestrina*) inferred from mitochondrial DNA. Am. J. Phys. Anthropol. 104:35–45.
- see dataset 53 and 54 and macaque datsets in Stelbrink et al. (2012)

Rutschmann F., Eriksson T., Schönenberger J., Conti E. 2004. Did Crypteroniaceae really disperse out of India? Molecular evidence from rbcL, ndhF, and rpl16 intron sequences. Int. J. Plant Sci. 165 Suppl:S69–S83.
- Gondwanan distribution, but origin of specimens uncertain
- authors: age of Asian lineages: Tertiary (Indian plate; dispersal in SE Asia)

Schaefer H., Heibl C., Renner S.S. 2009. Gourds afloat: a dated phylogeny reveals an Asian origin of the gourd family (Cucurbitaceae) and numerous oversea dispersal events. Proc. R. Soc. London B 276:843–851.
- high-level phylogeny
- authors: age of Asian lineages: Miocene-Paleocene

Schulte J.A., Melville J., Larson A. 2003. Molecular phylogenetic evidence for ancient divergence of lizard taxa on either side of Wallace's Line. Proc. R. Soc. London B 270:597–603.
- amphibulorine: only Australia and New Guinea; varanoid: Indonesian origin imprecise
- authors: age of Asian lineages: 120.5 / 150 Mya

Skinner A., Hugall A.F., Hutchinson M.N. 2011. Lygosomine phylogeny and the origins of Australian scincid lizards. J. Biogeogr. 38:1044–1058.
- SE Asia not well represented
- authors: age of Asian lineages: Eocene

Slechtová V., Bohlen J., Freyhof J., Ráb P. 2006. Molecular phylogeny of the Southeast Asian freshwater fish family Botiidae (Teleostei: Cobitoidea) and the origin of polyploidy in their evolution. Mol. Phylogenet. Evol. 39:529–541.
- geographic mixture among genera

Steppan S.J., Adkins R.M., Spinks P.Q., Hale C. 2005. Multigene phylogeny of the Old World mice, Murinae, reveals distinct geographic lineages and the declining utility of mitochondrial genes compared to nuclear genes. Mol. Phylogenet. Evol. 37:370–388.
- see dataset 58

Ueda S., Quek S.-P., Itioka T., Murase K., Itino T. 2010. Phylogeography of the *Coccus* scale insects inhabiting myrmecophytic *Macaranga* plants in Southeast Asia. Popul. Ecol. 52:137–146.
- distribution relatively small
- complex host relationship
- authors: age of Asian lineages: Pliocene

Vidal N., Marin J., Morini M., Donnellan S., Branch W.R., Thomas R., Vences M., Wynn A., Cruaud C., Blair Hedges S. 2010. Blindsnake evolutionary tree reveals long history on Gondwana. Biol. Lett. 6:558–561.

- excluded due to reviewer's comment: several taxa missing in dataset
- authors: age of Asian lineages: Miocene-Paleocene

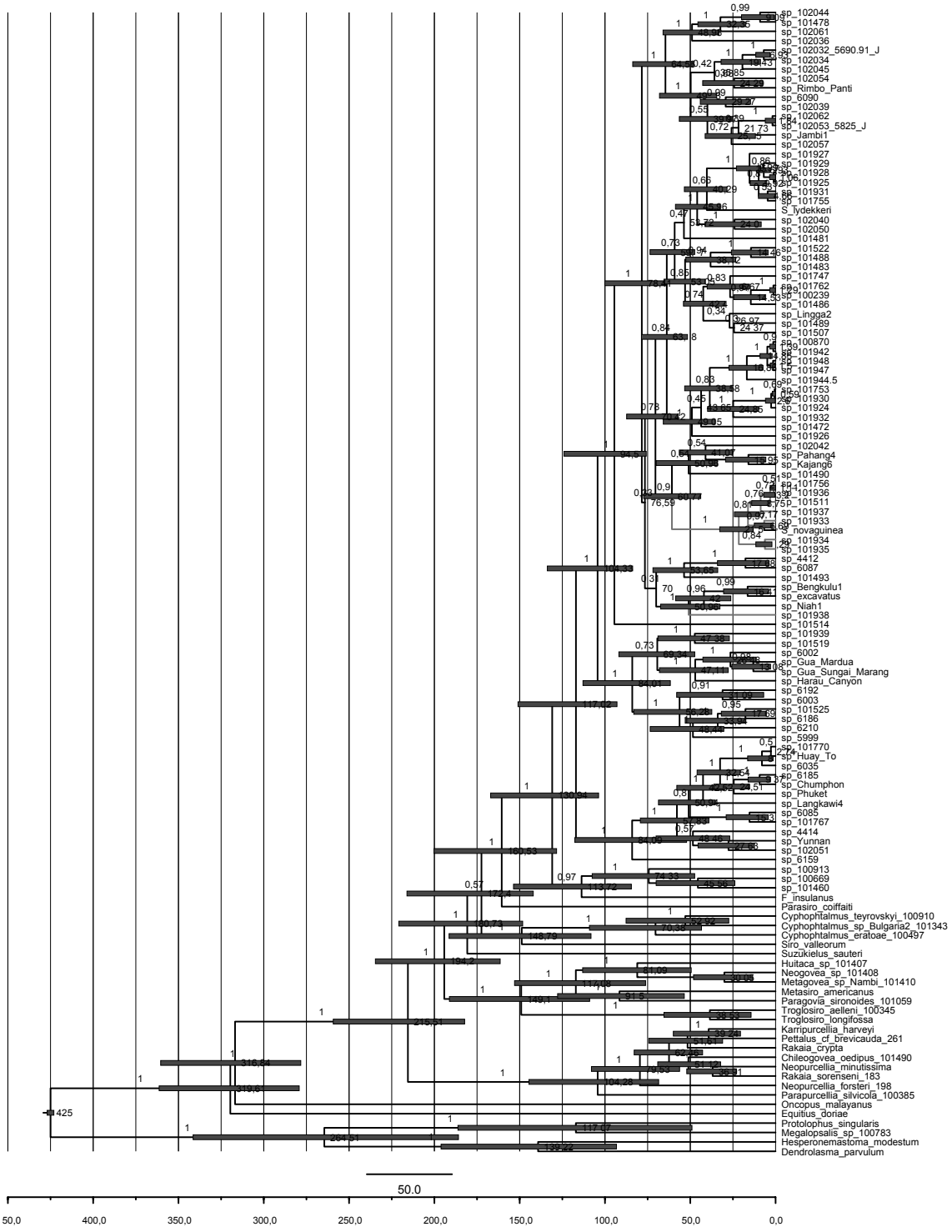
Wiens J.J., Sukumaran J., Pyron R.A., Brown R.M. 2009. Evolutionary and biogeographic origins of high tropical diversity in Old World frogs (Ranidae). Evolution 63:1217–1231.

- origin of speciments uncertain
- authors: age of Asian lineages: 87.4 Mya

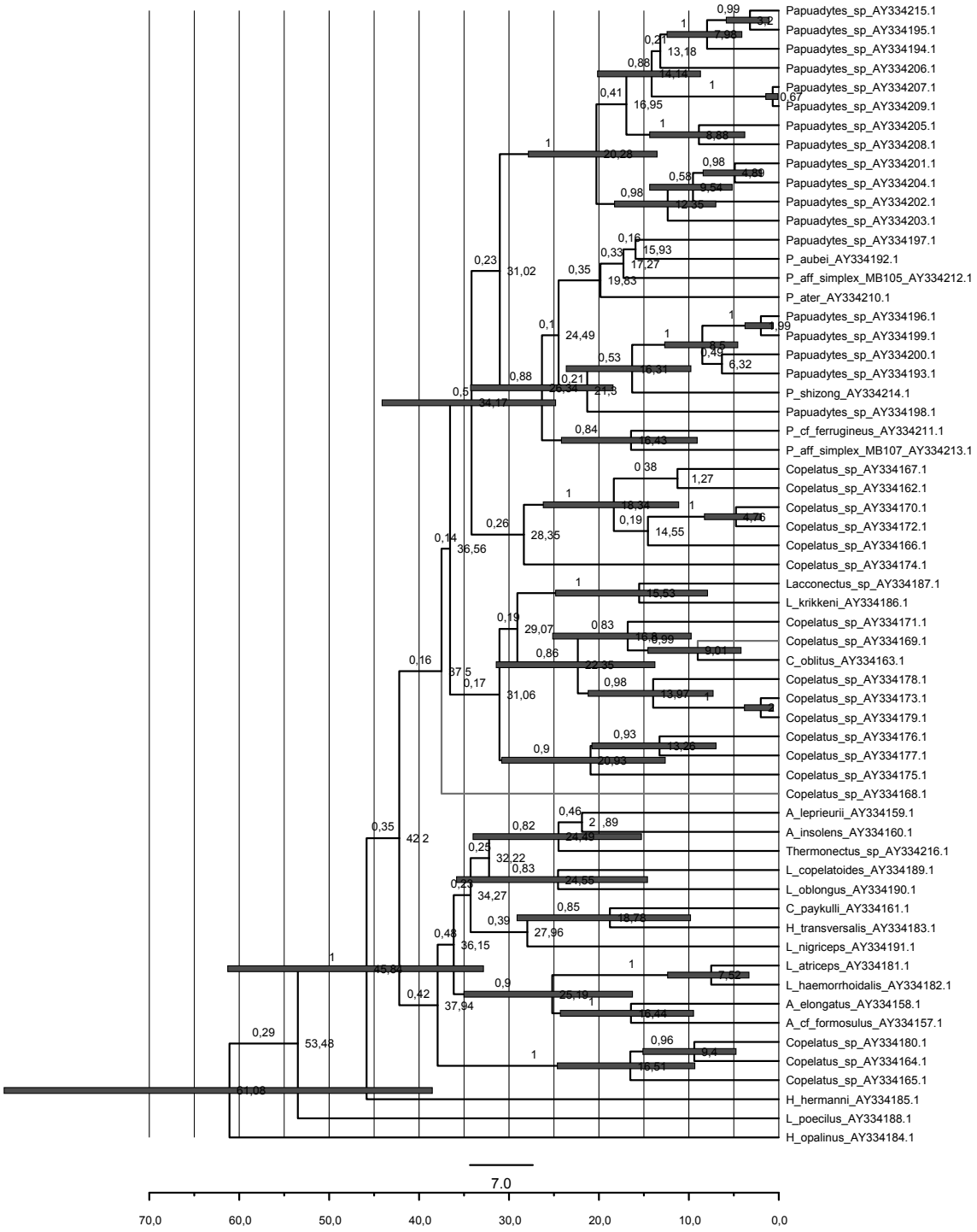
Zarowiecki M., Walton C., Torres E., McAlister E., Than Htun P., Sumrandee C., Sochanta T., Dinh T.H., Ching Ng L., Linton Y.-M. 2011. Pleistocene genetic connectivity in a widespread, open-habitat-adapted mosquito in the Indo-Oriental region. J. Biogeogr. 38: 1422-1432.

- population-based dataset;bad branch support
- see dataset 7

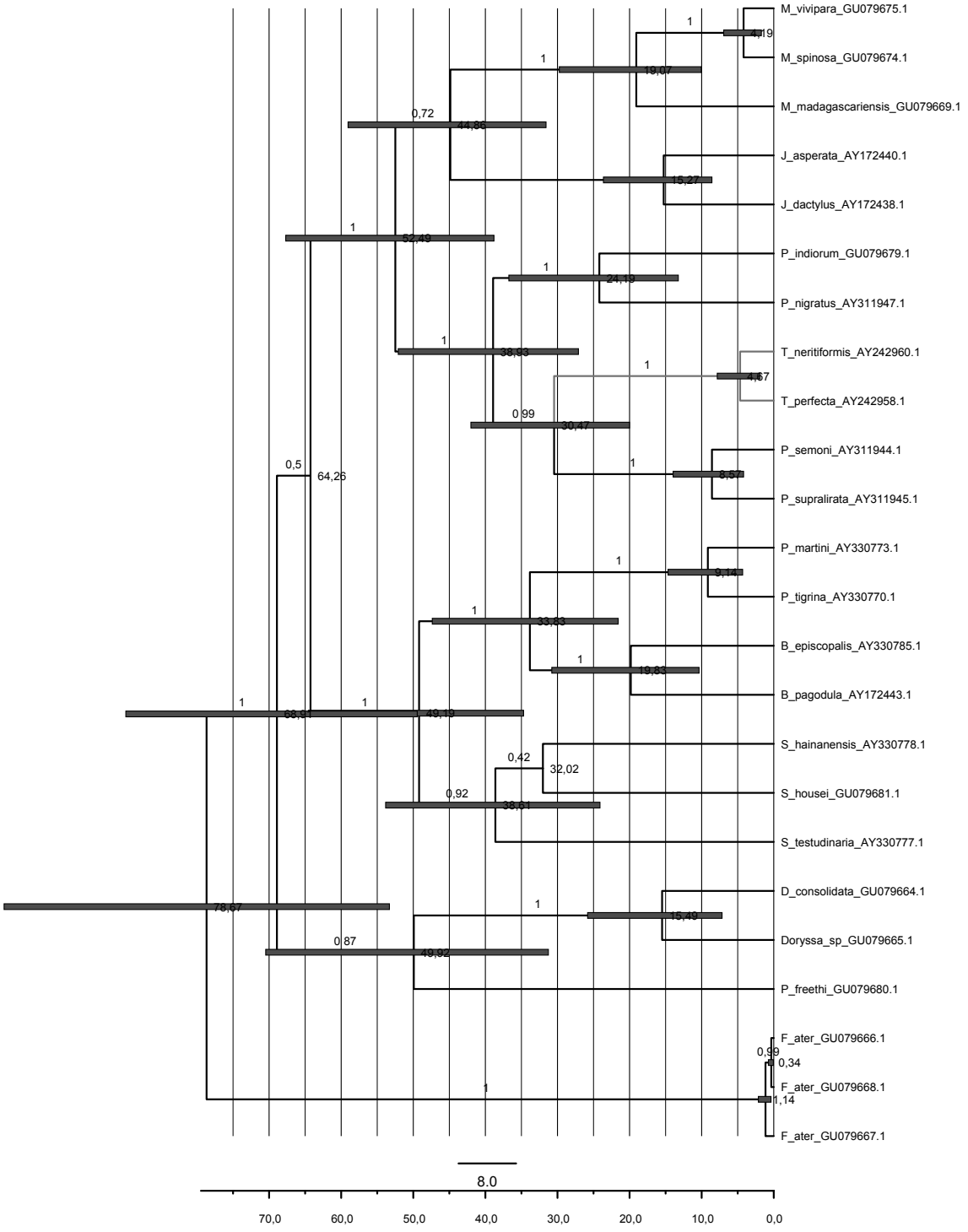
Supplementary Information – Chapter 2


sp_102044
sp_101478
sp_102061
sp_102036
sp_102032_5690.91_J
sp_102034
sp_102045
sp_102054
sp_Rimbo_Panti
sp_6090
sp_102039
sp_102062
sp_102053_5825_J
sp_Jambi1
sp_102057
sp_101927
sp_101929
sp_101928
sp_101925
sp_101931
sp_101755
S_Tydekkeri
sp_102040
sp_102050
sp_101481
sp_101522
sp_101488
sp_101483
sp_101747
sp_101762
sp_100239
sp_101486
sp_Lingga2
sp_101489
sp_101507
sp_100870
sp_101942
sp_101948
sp_101947
sp_101944.5
sp_101753
sp_101930
sp_101924
sp_101932
sp_101472
sp_101926
sp_102042
sp_Pahang4
sp_Kajang6
sp_101490
sp_101756
sp_101936
sp_101511
sp_101937
sp_101933
S_novaguinea
sp_101934
sp_101935
sp_4412
sp_6087
sp_101493
sp_Bengkulu1
sp_excavatus
sp_Niah1
sp_101938
sp_101514
sp_101939
sp_101519
sp_6002
sp_Gua_Mardua
sp_Gua_Sungai_Marang
sp_Harau_Canyon
sp_6192
sp_6003
sp_101525
sp_6186
sp_6210
sp_5999
sp_101770
sp_Huay_To
sp_6035
sp_6185
sp_Chumphon
sp_Phuket
sp_Langkawi4
sp_6085
sp_101767
sp_4414
sp_Yunnan
sp_102051
sp_6159
sp_100913
sp_100669
sp_101460
F_insulanus
Parasiro_coiffaiti
Cyphophtalmus_teyrovskyi_100910
Cyphophtalmus_sp_Bulgaria2_101343
Cyphophtalmus_eratoae_100497
Siro_valleorum
Suzukielus_sauteri
Huitaca_sp_101407
Neogovea_sp_101408
Metagovea_sp_Nambi_101410
Metasiro_americanus
Paragovia_sironoides_101059
Troglosiro_aelleni_100345
Troglosiro_longifossa
Karripurcellia_harveyi
Pettalus_cf_brevicauda_261
Rakaia_crypta
Chileogovea_oedipus_101490
Neopurcellia_minutissima
Rakaia_sorenseni_183
Neopurcellia_forsteri_198
Parapurcellia_silvicola_100385
Oncopus_malayanus
Equitius_doriae
Protolophus_singularis
Megalopsalis_sp_100783
Hesperonemastoma_modestum
Dendrolasma_parvulum
425
50.0
50,0
400,0
350,0
300,0
250,0
200,0
150,0
100,0
50,0
0,0

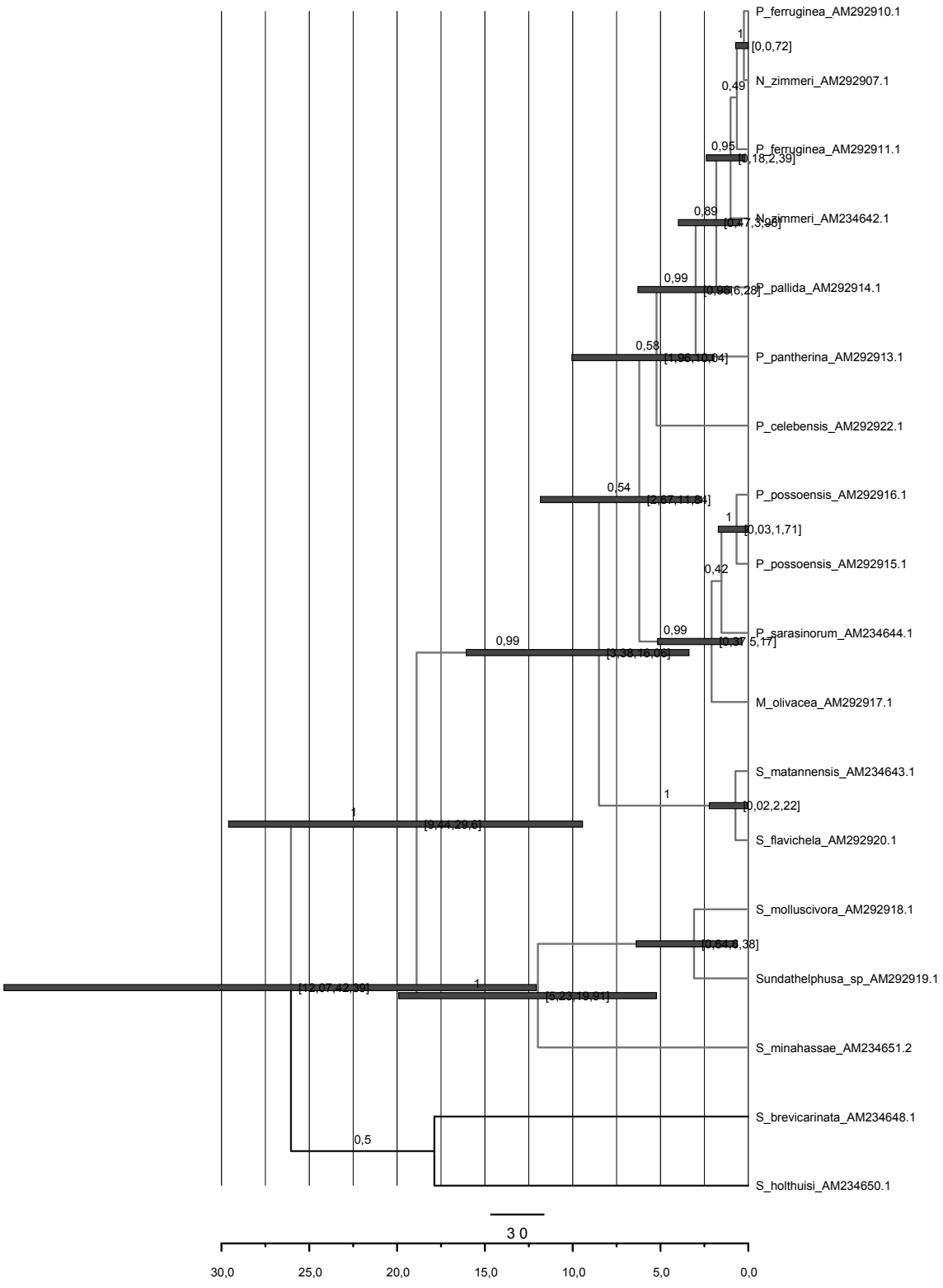
Dataset 1


Papuadytes_sp_AY334215.1
Papuadytes_sp_AY334195.1
Papuadytes_sp_AY334194.1
Papuadytes_sp_AY334206.1
Papuadytes_sp_AY334207.1
Papuadytes_sp_AY334209.1
Papuadytes_sp_AY334205.1
Papuadytes_sp_AY334208.1
Papuadytes_sp_AY334201.1
Papuadytes_sp_AY334204.1
Papuadytes_sp_AY334202.1
Papuadytes_sp_AY334203.1
Papuadytes_sp_AY334197.1
P_aubei_AY334192.1
P_aff_simplex_MB105_AY334212.1
P_ater_AY334210.1
Papuadytes_sp_AY334196.1
Papuadytes_sp_AY334199.1
Papuadytes_sp_AY334200.1
Papuadytes_sp_AY334193.1
P_shizong_AY334214.1
Papuadytes_sp_AY334198.1
P_cf_ferrugineus_AY334211.1
P_aff_simplex_MB107_AY334213.1
Copelatus_sp_AY334167.1
Copelatus_sp_AY334162.1
Copelatus_sp_AY334170.1
Copelatus_sp_AY334172.1
Copelatus_sp_AY334166.1
Copelatus_sp_AY334174.1
Lacconectus_sp_AY334187.1
L_krikkeni_AY334186.1
Copelatus_sp_AY334171.1
Copelatus_sp_AY334169.1
C_oblitus_AY334163.1
Copelatus_sp_AY334178.1
Copelatus_sp_AY334173.1
Copelatus_sp_AY334179.1
Copelatus_sp_AY334176.1
Copelatus_sp_AY334177.1
Copelatus_sp_AY334175.1
Copelatus_sp_AY334168.1
A_leprieurii_AY334159.1
A_insolens_AY334160.1
Thermonectus_sp_AY334216.1
L_copelatoides_AY334189.1
L_oblongus_AY334190.1
C_paykulli_AY334161.1
H_transversalis_AY334183.1
L_nigriceps_AY334191.1
L_atriceps_AY334181.1
L_haemorrhoidalis_AY334182.1
A_elongatus_AY334158.1
A_cf_formosulus_AY334157.1
Copelatus_sp_AY334180.1
Copelatus_sp_AY334164.1
Copelatus_sp_AY334165.1
H_hermanni_AY334185.1
L_poecilus_AY334188.1
H_opalinus_AY334184.1
7.0
70,0
60,0
50,0
40,0
30,0
20,0
10,0
0,0

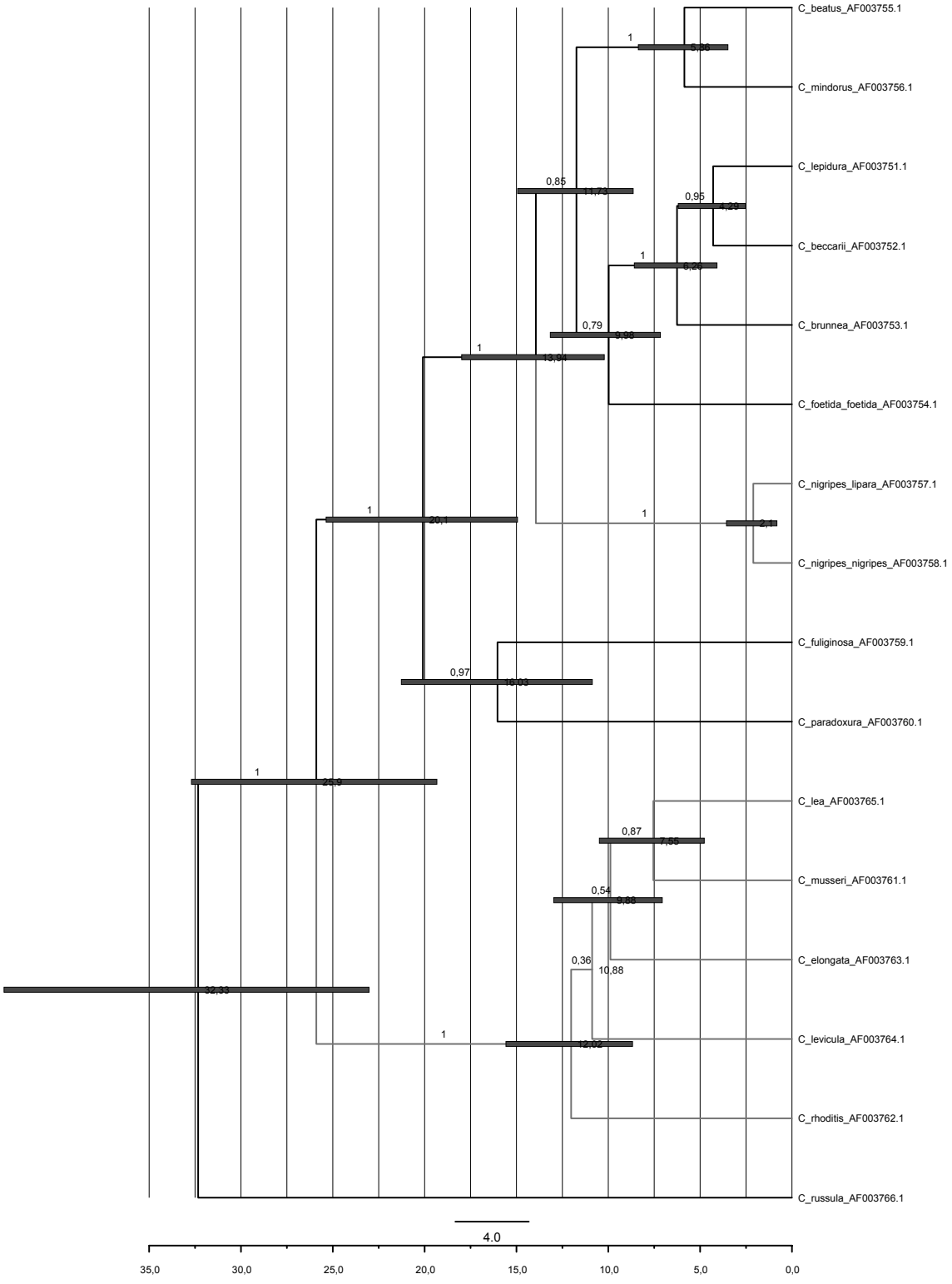
Dataset 2


M_vivipara_GU079675.1
M_spinosa_GU079674.1
M_madagascariensis_GU079669.1
J_asperata_AY172440.1
J_dactylus_AY172438.1
P_indiorum_GU079679.1
P_nigratus_AY311947.1
T_neritiformis_AY242960.1
T_perfecta_AY242958.1
P_semoni_AY311944.1
P_supralirata_AY311945.1
P_martini_AY330773.1
P_tigrina_AY330770.1
B_episcopalis_AY330785.1
B_pagodula_AY172443.1
S_hainanensis_AY330778.1
S_housei_GU079681.1
S_testudinaria_AY330777.1
D_consolidata_GU079664.1
Doryssa_sp_GU079665.1
P_freethi_GU079680.1
F_ater_GU079666.1
F_ater_GU079668.1
F_ater_GU079667.1
8.0
70,0
60,0
50,0
40,0
30,0
20,0
10,0
0,0

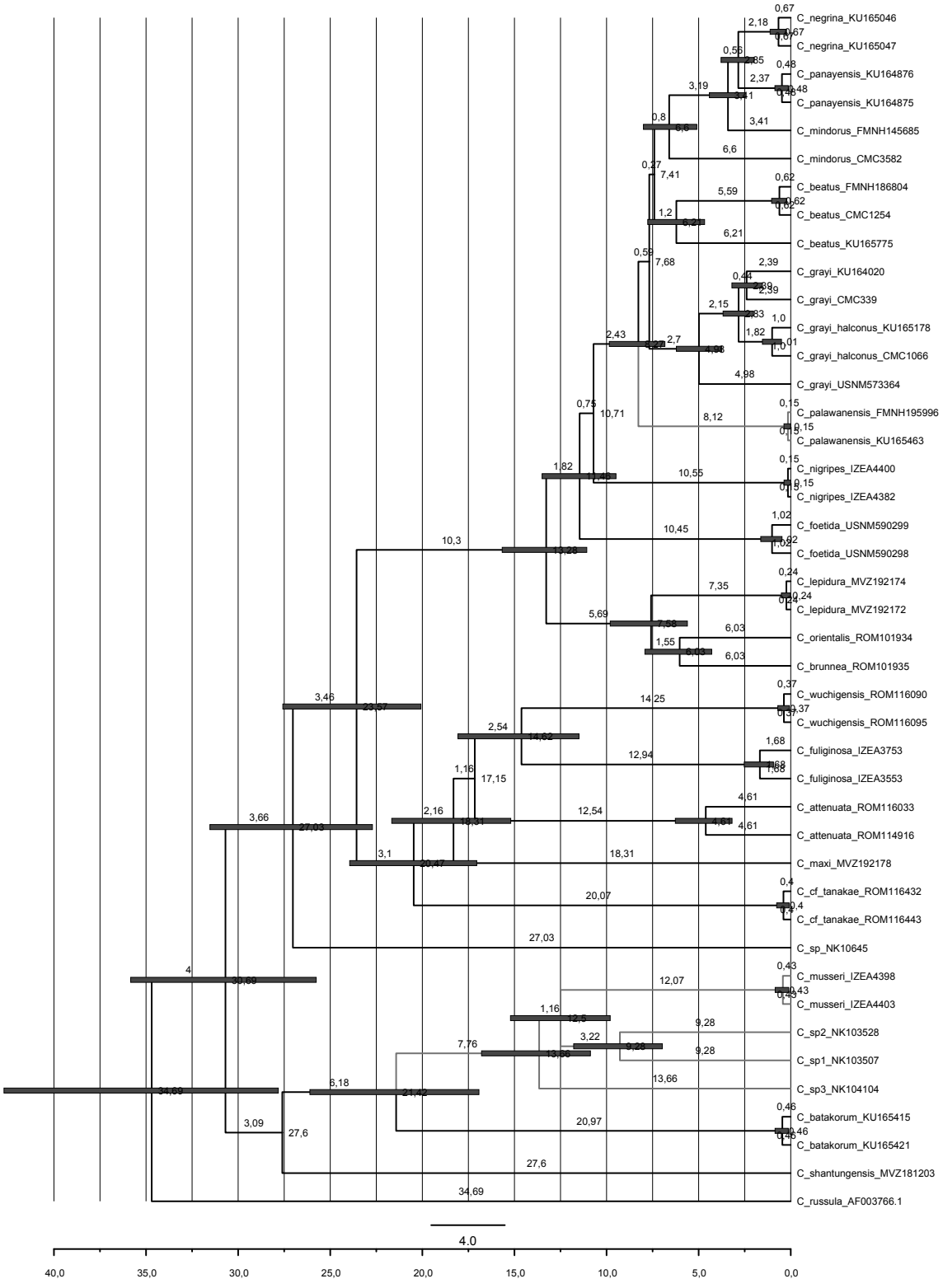
Dataset 3


P_ferruginea_AM292910.1
1
[0,0,72]
N_zimmeri_AM292907.1
0,49
0,95
P_ferruginea_AM292911.1
[0,18,2,39]
0,89
N_zimmeri_AM234642.1
0,99
P_pallida_AM292914.1
0,58
P_pantherina_AM292913.1
P_celebensis_AM292922.1
0,54
P_possoensis_AM292916.1
1
[0,03,1,71]
P_possoensis_AM292915.1
0,42
0,99
P_sarasinorum_AM234644.1
0,99
M_olivacea_AM292917.1
S_matannensis_AM234643.1
1
[0,02,2,22]
1
[9,44,29,6]
S_flavichela_AM292920.1
S_molluscivora_AM292918.1
Sundathelphusa_sp_AM292919.1
[12,07,42,39]
1
[5,23,19,91]
S_minahassae_AM234651.2
S_brevicarinata_AM234648.1
0,5
S_holthuisi_AM234650.1
30
30,0
25,0
20,0
15,0
10,0
5,0
0,0

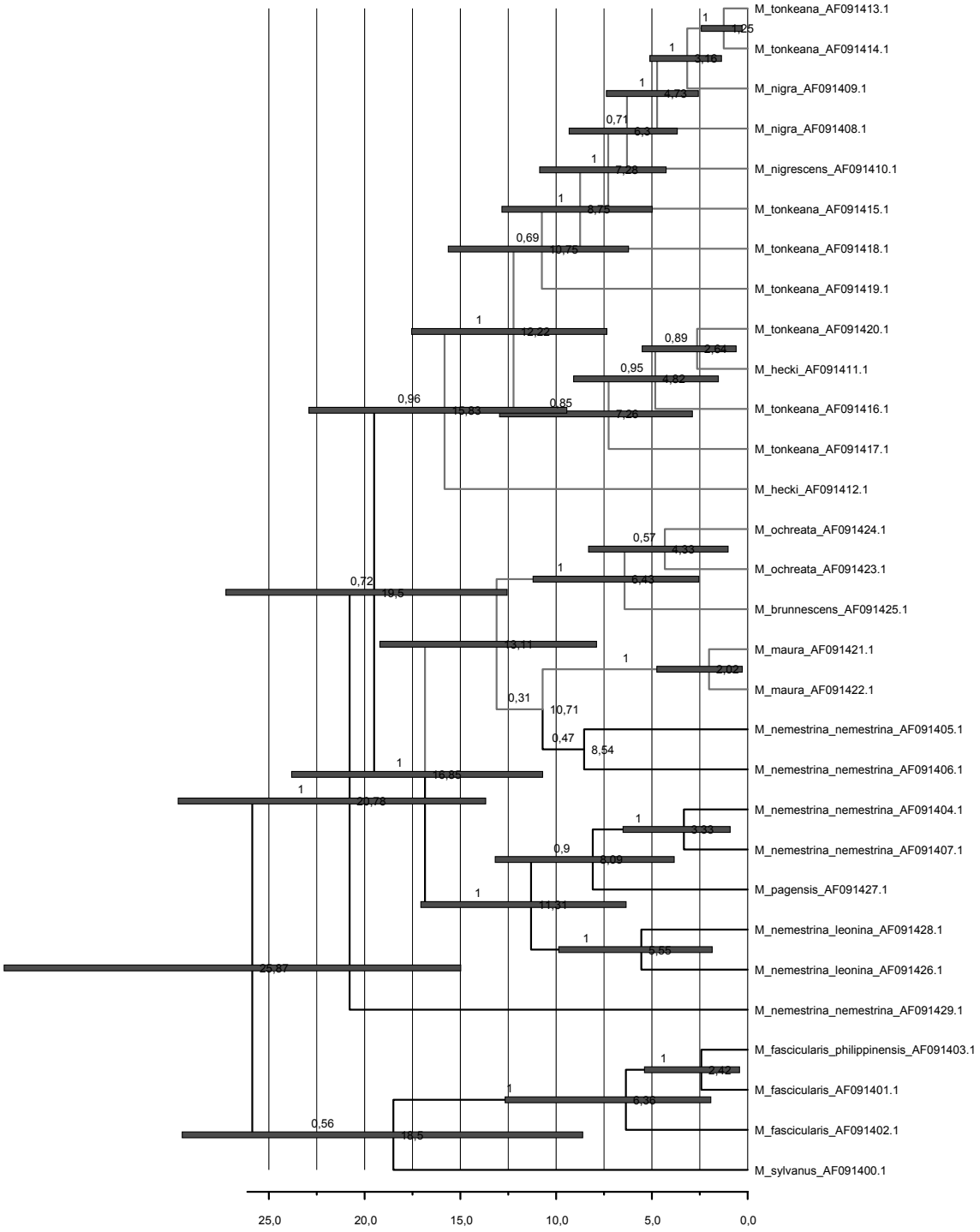
Dataset 4


C_beatus_AF003755.1
C_mindorus_AF003756.1
C_lepidura_AF003751.1
C_beccarii_AF003752.1
C_brunnea_AF003753.1
C_foetida_foetida_AF003754.1
C_nigripes_lipara_AF003757.1
C_nigripes_nigripes_AF003758.1
C_fuliginosa_AF003759.1
C_paradoxura_AF003760.1
C_lea_AF003765.1
C_musseri_AF003761.1
C_elongata_AF003763.1
C_levicula_AF003764.1
C_rhoditis_AF003762.1
C_russula_AF003766.1
1
0,85
11,73
0,95
1
0,79
1
13,94
1
20,1
1
0,97
1
0,87
0,54
0,36
10,88
1
4.0
35,0
30,0
25,0
20,0
15,0
10,0
5,0
0,0

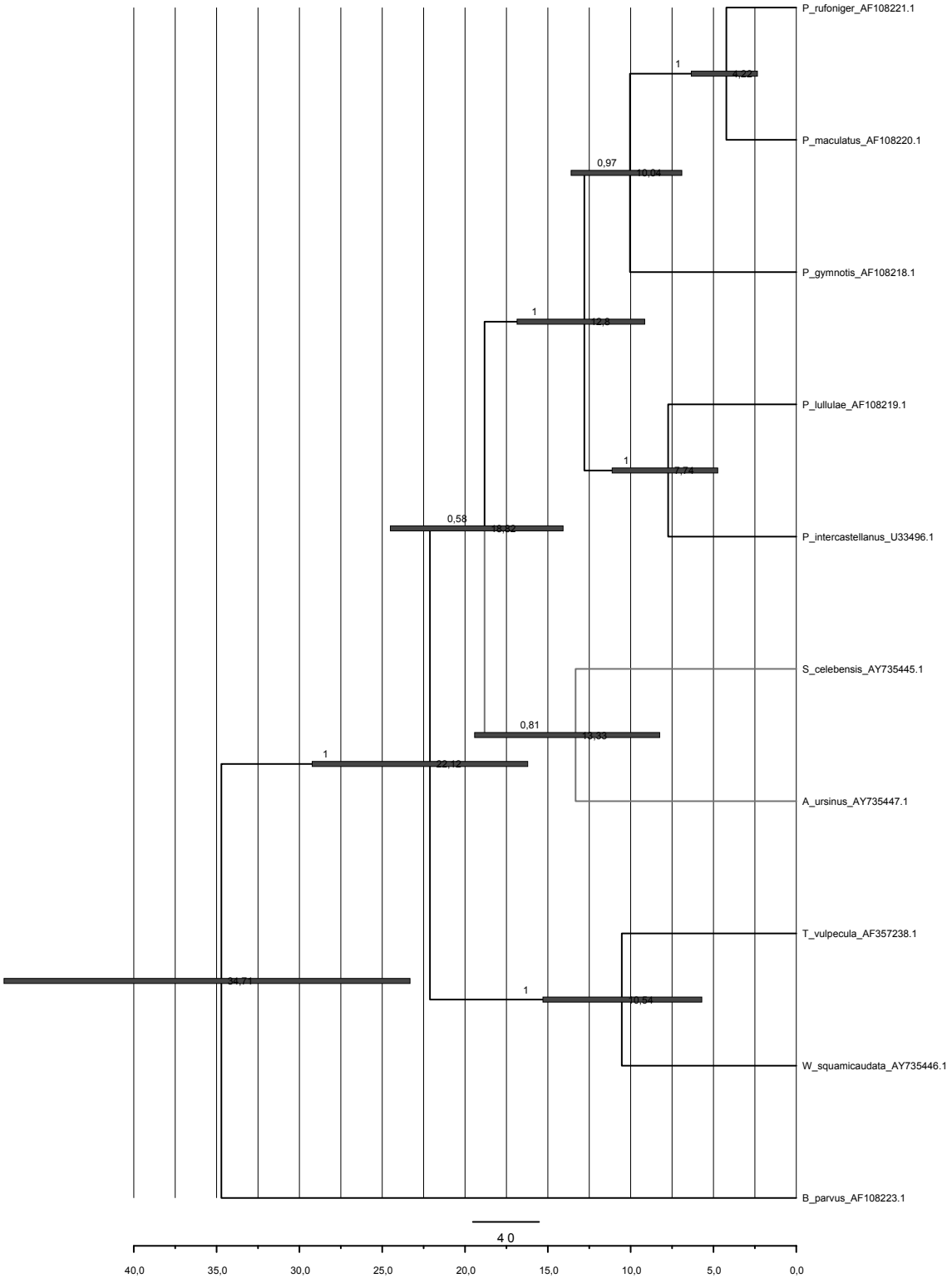
Dataset 5


C_negrina_KU165046
C_negrina_KU165047
C_panayensis_KU164876
C_panayensis_KU164875
C_mindorus_FMNH145685
C_mindorus_CMC3582
C_beatus_FMNH186804
C_beatus_CMC1254
C_beatus_KU165775
C_grayi_KU164020
C_grayi_CMC339
C_grayi_halconus_KU165178
C_grayi_halconus_CMC1066
C_grayi_USNM573364
C_palawanensis_FMNH195996
C_palawanensis_KU165463
C_nigripes_IZEA4400
C_nigripes_IZEA4382
C_foetida_USNM590299
C_foetida_USNM590298
C_lepidura_MVZ192174
C_lepidura_MVZ192172
C_orientalis_ROM101934
C_brunnea_ROM101935
C_wuchigensis_ROM116090
C_wuchigensis_ROM116095
C_fuliginosa_IZEA3753
C_fuliginosa_IZEA3553
C_attenuata_ROM116033
C_attenuata_ROM114916
C_maxi_MVZ192178
C_cf_tanakae_ROM116432
C_cf_tanakae_ROM116443
C_sp_NK10645
C_musseri_IZEA4398
C_musseri_IZEA4403
C_sp2_NK103528
C_sp1_NK103507
C_sp3_NK104104
C_batakorum_KU165415
C_batakorum_KU165421
C_shantungensis_MVZ181203
C_russula_AF003766.1
4.0
40,0
35,0
30,0
25,0
20,0
15,0
10,0
5,0
0,0

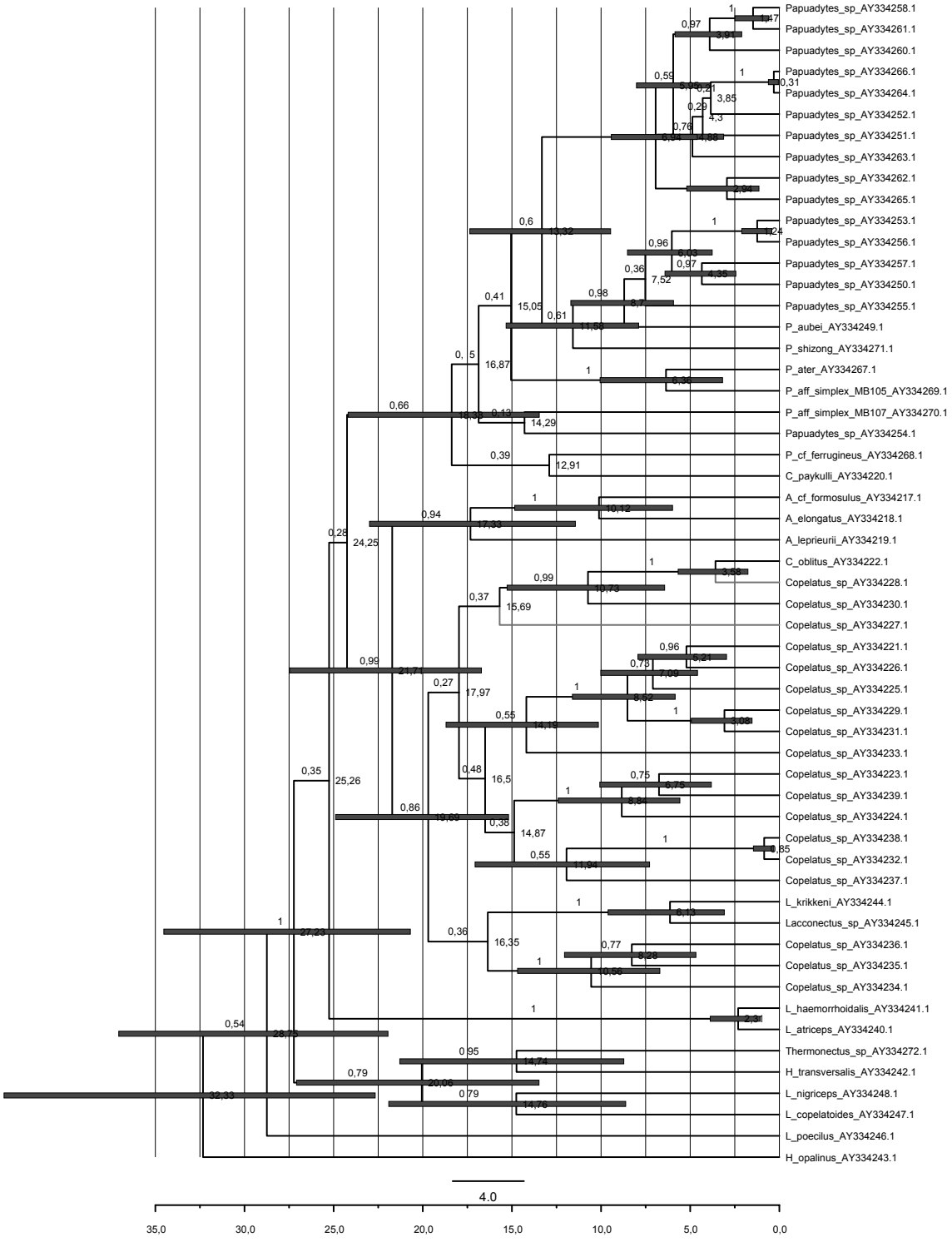
Dataset 6


M_tonkeana_AF091413.1
1
1,25
M_tonkeana_AF091414.1
1
3,16
M_nigra_AF091409.1
1
4,73
M_nigra_AF091408.1
0,71
6,3
M_nigrescens_AF091410.1
1
7,26
M_tonkeana_AF091415.1
1
8,75
M_tonkeana_AF091418.1
0,69
10,75
M_tonkeana_AF091419.1
M_tonkeana_AF091420.1
1
12,22
M_hecki_AF091411.1
0,89
2,64
M_tonkeana_AF091416.1
0,95
4,82
M_tonkeana_AF091417.1
0,96
15,63
0,85
7,26
M_hecki_AF091412.1
M_ochreata_AF091424.1
0,57
4,33
M_ochreata_AF091423.1
1
6,43
M_brunnescens_AF091425.1
0,72
19,5
M_maura_AF091421.1
13,11
1
2,02
M_maura_AF091422.1
0,31
10,71
M_nemestrina_nemestrina_AF091405.1
0,47
8,54
M_nemestrina_nemestrina_AF091406.1
1
16,85
M_nemestrina_nemestrina_AF091404.1
1
20,79
1
3,33
M_nemestrina_nemestrina_AF091407.1
0,9
6,09
M_pagensis_AF091427.1
1
11,31
M_nemestrina_leonina_AF091428.1
1
0,55
M_nemestrina_leonina_AF091426.1
25,87
M_nemestrina_nemestrina_AF091429.1
M_fascicularis_philippinensis_AF091403.1
1
2,42
M_fascicularis_AF091401.1
1
6,36
M_fascicularis_AF091402.1
0,56
18,5
M_sylvanus_AF091400.1
25,0
20,0
15,0
10,0
5,0
0,0

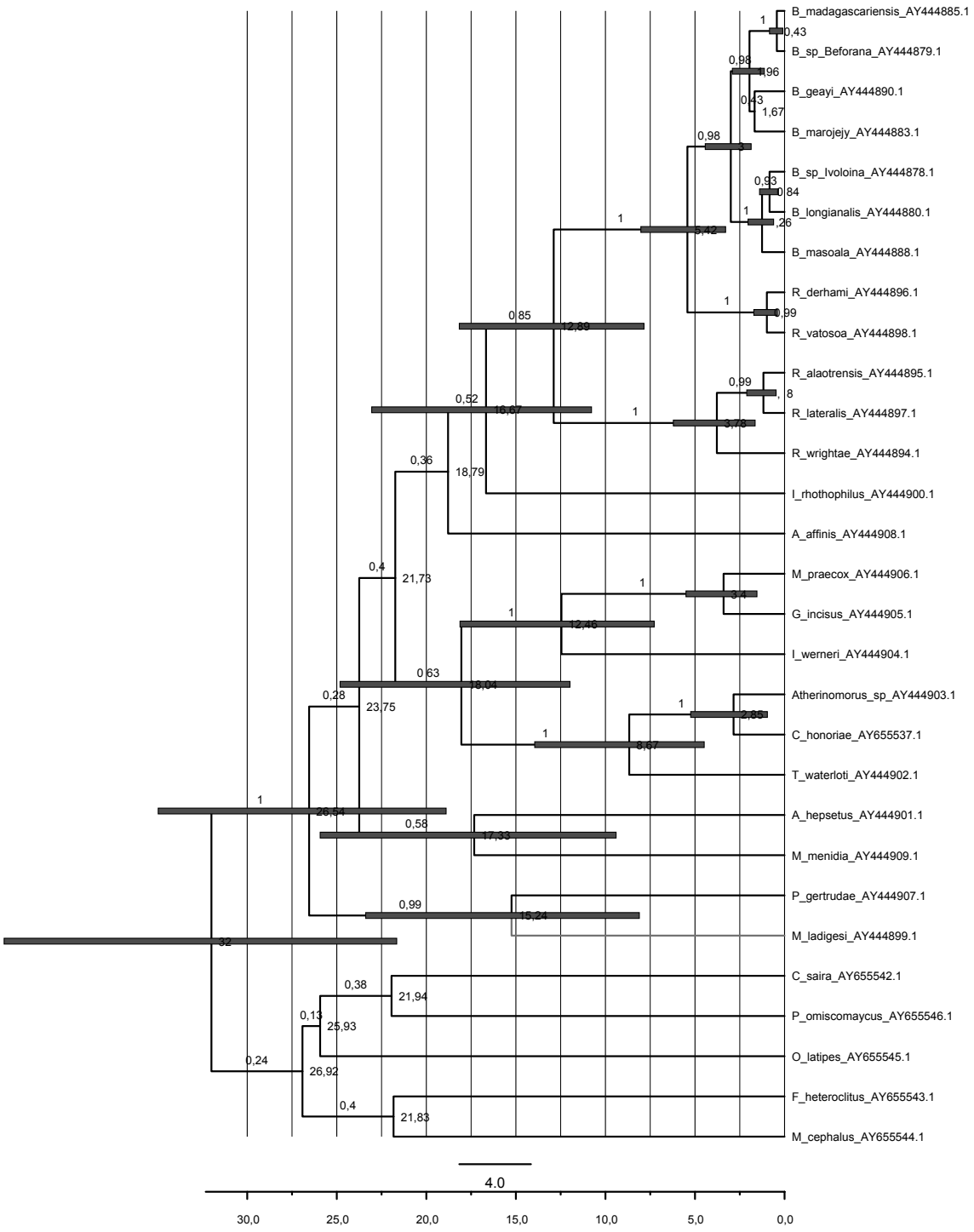
Dataset 7


P_rufoniger_AF108221.1
1
1,22
P_maculatus_AF108220.1
0,97
10,04
P_gymnotis_AF108218.1
1
12,8
P_lullulae_AF108219.1
1
7,74
0,58
15,12
P_intercastellanus_U33496.1
S_celebensis_AY735445.1
0,81
13,33
1
22,12
A_ursinus_AY735447.1
T_vulpecula_AF357238.1
34,71
1
0,64
W_squamicaudata_AY735446.1
B_parvus_AF108223.1
40
40,0
35,0
30,0
25,0
20,0
15,0
10,0
5,0
0,0

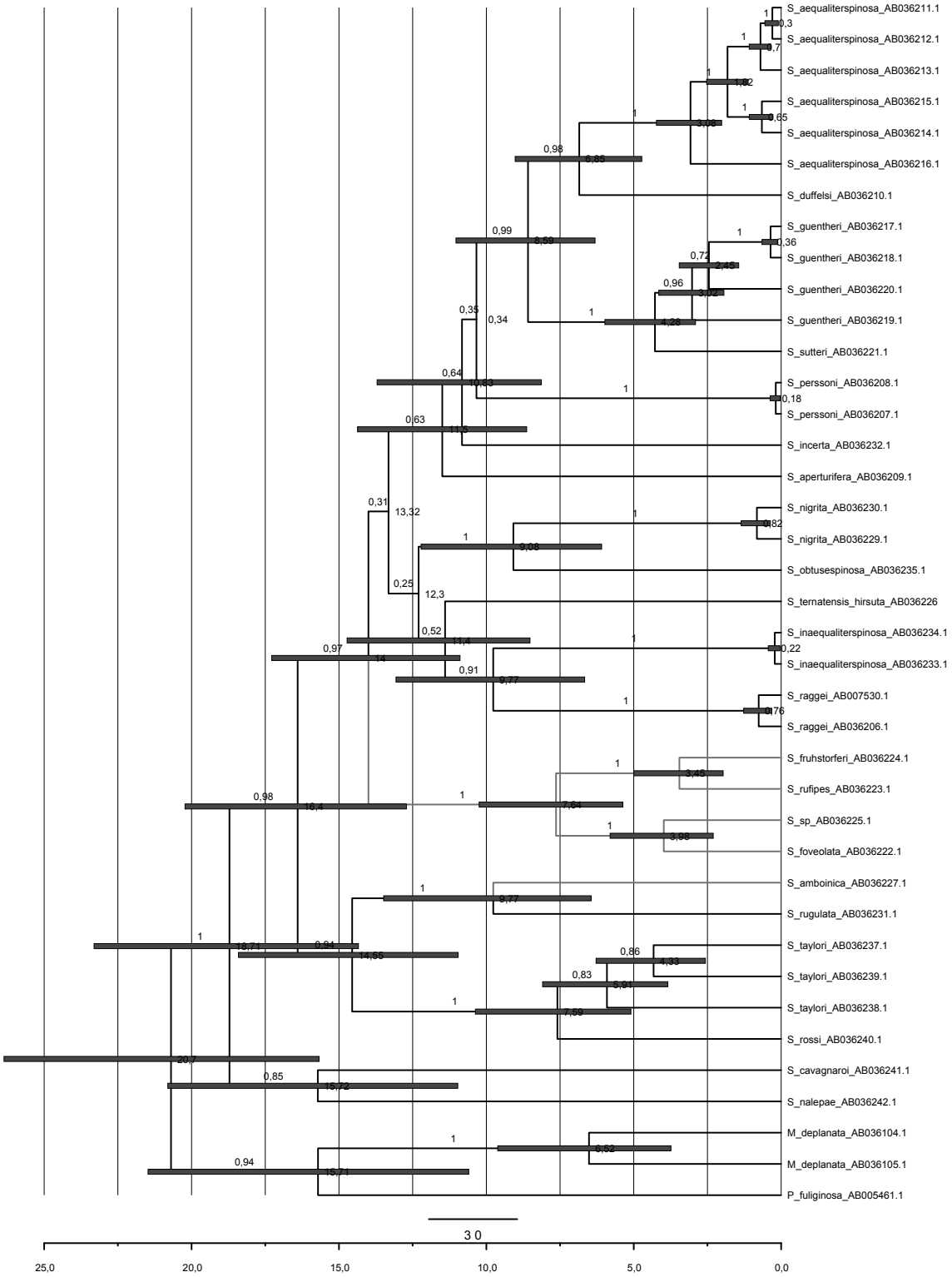
Dataset 8


Papuadytes_sp_AY334258.1
Papuadytes_sp_AY334261.1
Papuadytes_sp_AY334260.1
Papuadytes_sp_AY334266.1
Papuadytes_sp_AY334264.1
Papuadytes_sp_AY334252.1
Papuadytes_sp_AY334251.1
Papuadytes_sp_AY334263.1
Papuadytes_sp_AY334262.1
Papuadytes_sp_AY334265.1
Papuadytes_sp_AY334253.1
Papuadytes_sp_AY334256.1
Papuadytes_sp_AY334257.1
Papuadytes_sp_AY334250.1
Papuadytes_sp_AY334255.1
P_aubei_AY334249.1
P_shizong_AY334271.1
P_ater_AY334267.1
P_aff_simplex_MB105_AY334269.1
P_aff_simplex_MB107_AY334270.1
Papuadytes_sp_AY334254.1
P_cf_ferrugineus_AY334268.1
C_paykulli_AY334220.1
A_cf_formosulus_AY334217.1
A_elongatus_AY334218.1
A_leprieurii_AY334219.1
C_oblitus_AY334222.1
Copelatus_sp_AY334228.1
Copelatus_sp_AY334230.1
Copelatus_sp_AY334227.1
Copelatus_sp_AY334221.1
Copelatus_sp_AY334226.1
Copelatus_sp_AY334225.1
Copelatus_sp_AY334229.1
Copelatus_sp_AY334231.1
Copelatus_sp_AY334233.1
Copelatus_sp_AY334223.1
Copelatus_sp_AY334239.1
Copelatus_sp_AY334224.1
Copelatus_sp_AY334238.1
Copelatus_sp_AY334232.1
Copelatus_sp_AY334237.1
L_krikkeni_AY334244.1
Lacconectus_sp_AY334245.1
Copelatus_sp_AY334236.1
Copelatus_sp_AY334235.1
Copelatus_sp_AY334234.1
L_haemorrhoidalis_AY334241.1
L_atriceps_AY334240.1
Thermonectus_sp_AY334272.1
H_transversalis_AY334242.1
L_nigriceps_AY334248.1
L_copelatoides_AY334247.1
L_poecilus_AY334246.1
H_opalinus_AY334243.1
4.0
35,0
30,0
25,0
20,0
15,0
10,0
5,0
0,0

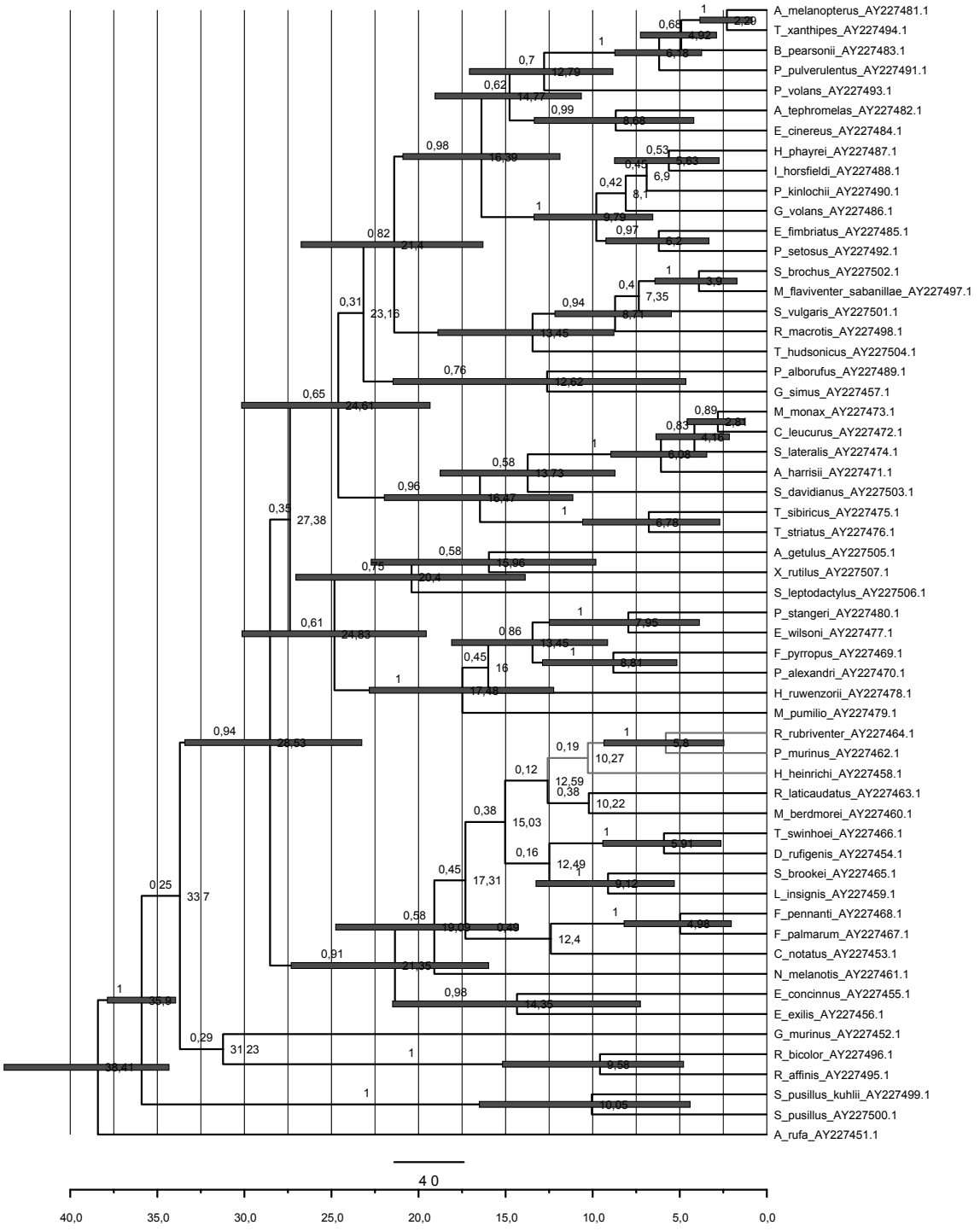
Dataset 9


B_madagascariensis_AY444885.1
B_sp_Beforana_AY444879.1
B_geayi_AY444890.1
B_marojejy_AY444883.1
B_sp_Ivoloina_AY444878.1
B_longianalis_AY444880.1
B_masoala_AY444888.1
R_derhami_AY444896.1
R_vatosoa_AY444898.1
R_alaotrensis_AY444895.1
R_lateralis_AY444897.1
R_wrightae_AY444894.1
I_rhothophilus_AY444900.1
A_affinis_AY444908.1
M_praecox_AY444906.1
G_incisus_AY444905.1
I_werneri_AY444904.1
Atherinomorus_sp_AY444903.1
C_honoriae_AY655537.1
T_waterloti_AY444902.1
A_hepsetus_AY444901.1
M_menidia_AY444909.1
P_gertrudae_AY444907.1
M_ladigesi_AY444899.1
C_saira_AY655542.1
P_omiscomaycus_AY655546.1
O_latipes_AY655545.1
F_heteroclitus_AY655543.1
M_cephalus_AY655544.1
4.0
30,0
25,0
20,0
15,0
10,0
5,0
0,0

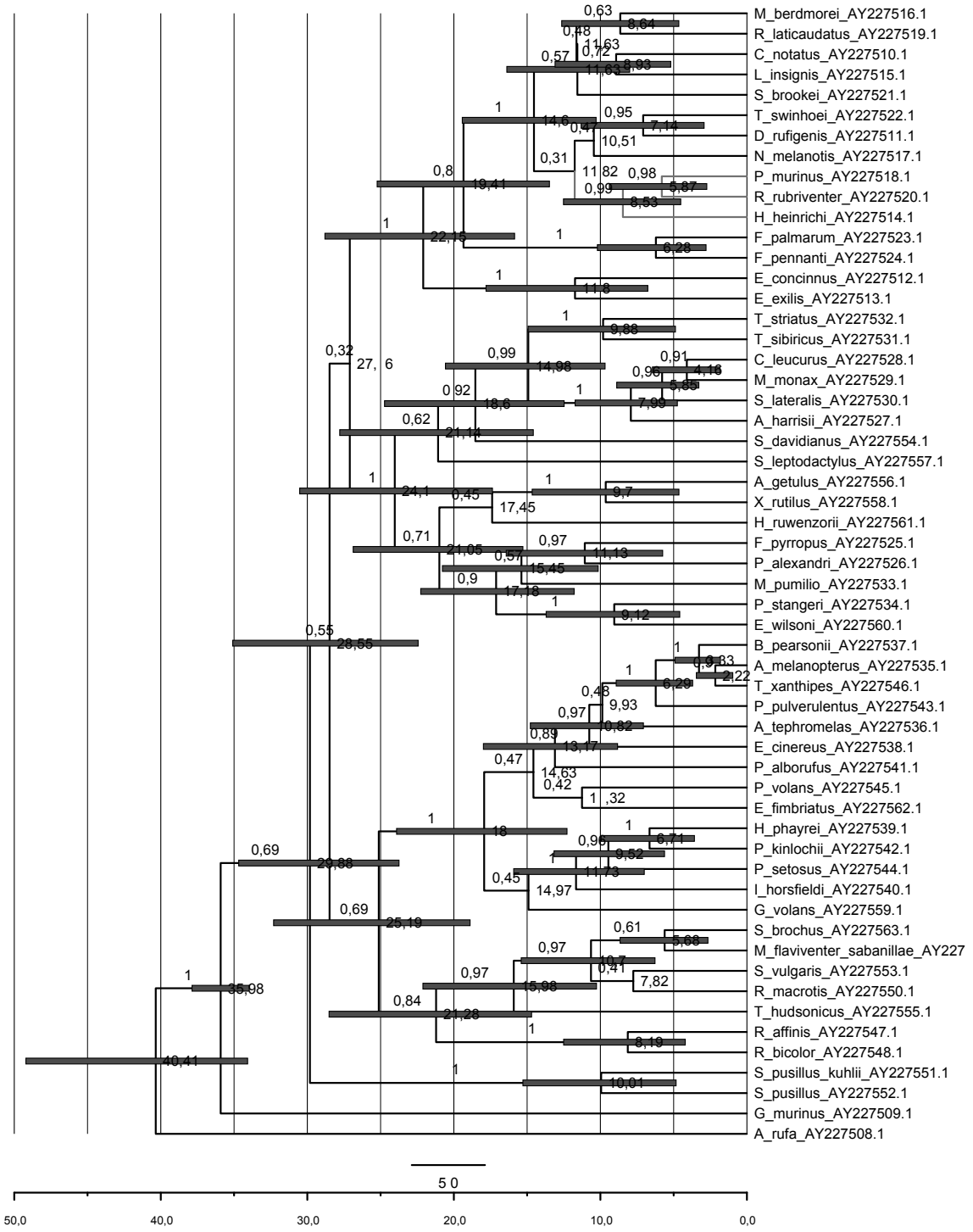
Dataset 10


S_aequaliterspinosa_AB036211.1
S_aequaliterspinosa_AB036212.1
S_aequaliterspinosa_AB036213.1
S_aequaliterspinosa_AB036215.1
S_aequaliterspinosa_AB036214.1
S_aequaliterspinosa_AB036216.1
S_duffelsi_AB036210.1
S_guentheri_AB036217.1
S_guentheri_AB036218.1
S_guentheri_AB036220.1
S_guentheri_AB036219.1
S_sutteri_AB036221.1
S_perssoni_AB036208.1
S_perssoni_AB036207.1
S_incerta_AB036232.1
S_aperturifera_AB036209.1
S_nigrita_AB036230.1
S_nigrita_AB036229.1
S_obtusespinosa_AB036235.1
S_ternatensis_hirsuta_AB036226
S_inaequaliterspinosa_AB036234.1
S_inaequaliterspinosa_AB036233.1
S_raggei_AB007530.1
S_raggei_AB036206.1
S_fruhstorferi_AB036224.1
S_rufipes_AB036223.1
S_sp_AB036225.1
S_foveolata_AB036222.1
S_amboinica_AB036227.1
S_rugulata_AB036231.1
S_taylori_AB036237.1
S_taylori_AB036239.1
S_taylori_AB036238.1
S_rossi_AB036240.1
S_cavagnaroi_AB036241.1
S_nalepae_AB036242.1
M_deplanata_AB036104.1
M_deplanata_AB036105.1
P_fuliginosa_AB005461.1
3 0
25,0
20,0
15,0
10,0
5,0
0,0

Dataset 11


A_melanopterus_AY227481.1
T_xanthipes_AY227494.1
B_pearsonii_AY227483.1
P_pulverulentus_AY227491.1
P_volans_AY227493.1
A_tephromelas_AY227482.1
E_cinereus_AY227484.1
H_phayrei_AY227487.1
I_horsfieldi_AY227488.1
P_kinlochii_AY227490.1
G_volans_AY227486.1
E_fimbriatus_AY227485.1
P_setosus_AY227492.1
S_brochus_AY227502.1
M_flaviventer_sabanillae_AY227497.1
S_vulgaris_AY227501.1
R_macrotis_AY227498.1
T_hudsonicus_AY227504.1
P_alborufus_AY227489.1
G_simus_AY227457.1
M_monax_AY227473.1
C_leucurus_AY227472.1
S_lateralis_AY227474.1
A_harrisii_AY227471.1
S_davidianus_AY227503.1
T_sibiricus_AY227475.1
T_striatus_AY227476.1
A_getulus_AY227505.1
X_rutilus_AY227507.1
S_leptodactylus_AY227506.1
P_stangeri_AY227480.1
E_wilsoni_AY227477.1
F_pyrropus_AY227469.1
P_alexandri_AY227470.1
H_ruwenzorii_AY227478.1
M_pumilio_AY227479.1
R_rubriventer_AY227464.1
P_murinus_AY227462.1
H_heinrichi_AY227458.1
R_laticaudatus_AY227463.1
M_berdmorei_AY227460.1
T_swinhoei_AY227466.1
D_rufigenis_AY227454.1
S_brookei_AY227465.1
L_insignis_AY227459.1
F_pennanti_AY227468.1
F_palmarum_AY227467.1
C_notatus_AY227453.1
N_melanotis_AY227461.1
E_concinnus_AY227455.1
E_exilis_AY227456.1
G_murinus_AY227452.1
R_bicolor_AY227496.1
R_affinis_AY227495.1
S_pusillus_kuhlii_AY227499.1
S_pusillus_AY227500.1
A_rufa_AY227451.1
40
40,0
35,0
30,0
25,0
20,0
15,0
10,0
5,0
0,0

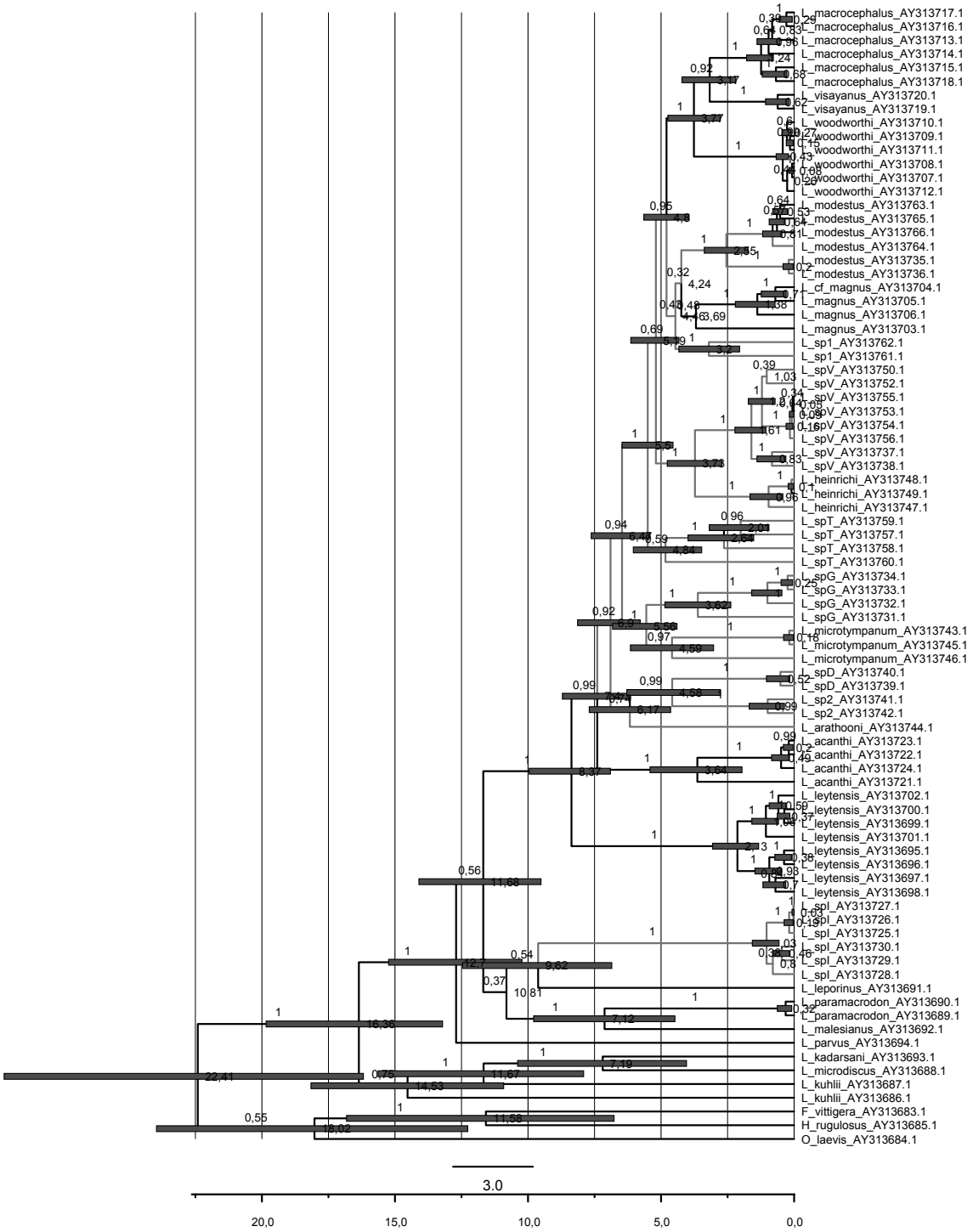
Dataset 12


M_berdmorei_AY227516.1
R_laticaudatus_AY227519.1
C_notatus_AY227510.1
L_insignis_AY227515.1
S_brookei_AY227521.1
T_swinhoei_AY227522.1
D_rufigenis_AY227511.1
N_melanotis_AY227517.1
P_murinus_AY227518.1
R_rubriventer_AY227520.1
H_heinrichi_AY227514.1
F_palmarum_AY227523.1
F_pennanti_AY227524.1
E_concinnus_AY227512.1
E_exilis_AY227513.1
T_striatus_AY227532.1
T_sibiricus_AY227531.1
C_leucurus_AY227528.1
M_monax_AY227529.1
S_lateralis_AY227530.1
A_harrisii_AY227527.1
S_davidianus_AY227554.1
S_leptodactylus_AY227557.1
A_getulus_AY227556.1
X_rutilus_AY227558.1
H_ruwenzorii_AY227561.1
F_pyrropus_AY227525.1
P_alexandri_AY227526.1
M_pumilio_AY227533.1
P_stangeri_AY227534.1
E_wilsoni_AY227560.1
B_pearsonii_AY227537.1
A_melanopterus_AY227535.1
T_xanthipes_AY227546.1
P_pulverulentus_AY227543.1
A_tephromelas_AY227536.1
E_cinereus_AY227538.1
P_alborufus_AY227541.1
P_volans_AY227545.1
E_fimbriatus_AY227562.1
H_phayrei_AY227539.1
P_kinlochii_AY227542.1
P_setosus_AY227544.1
I_horsfieldi_AY227540.1
G_volans_AY227559.1
S_brochus_AY227563.1
M_flaviventer_sabanillae_AY227
S_vulgaris_AY227553.1
R_macrotis_AY227550.1
T_hudsonicus_AY227555.1
R_affinis_AY227547.1
R_bicolor_AY227548.1
S_pusillus_kuhlii_AY227551.1
S_pusillus_AY227552.1
G_murinus_AY227509.1
A_rufa_AY227508.1
5 0
50,0
40,0
30,0
20,0
10,0
0,0

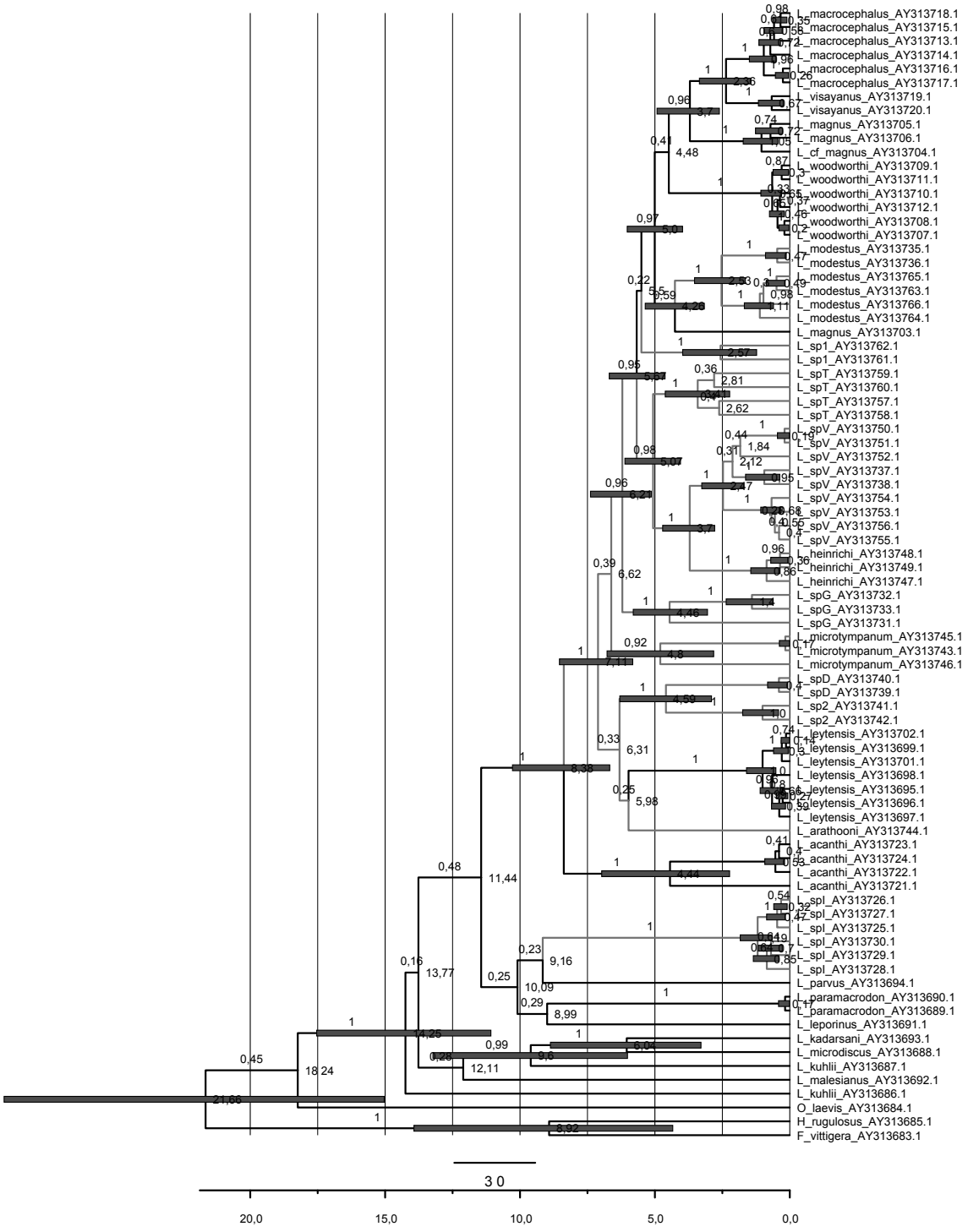
Dataset 13


M_forstenii_AF394629.1
0,08
M_forstenii_AF394630.1
0,42
M_freycinet_freycinet_AF394631.1
0,1
M_freycinet_freycinet_AF394632.1
M_decollatus_AF394625.1
M_freycinet_quoyii_AF394628.1
M_eremita_AF394627.1
0,12
M_eremita_AF394626.1
M_reinwardt_AF394633.1
M_layardi_AF394635.1
0,29
M_layardi_AF394634.1
M_pritchardii_AF394636.1
M_cumingii_AF394624.1
M_tenimberensis_AF394637.1
E_wallacei_AF394623.1
0,08
E_wallacei_AF394622.1
A_lathami_AF394616.1
0,15
A_lathami_AF394615.1
A_arfakianus_AF394617.1
T_fuscirostris_AF394620.1
L_ocellata_AF394619.1
0,08
L_ocellata_AF394618.1
M_maleo_AF394621.1
O_vetula_AF394614.1
D_obscurus_AF394611.1
P_cristatus_AF394612.1
N_meleagris_AF394613.1
3.0
25,0
20,0
15,0
10,0
5,0
0,0

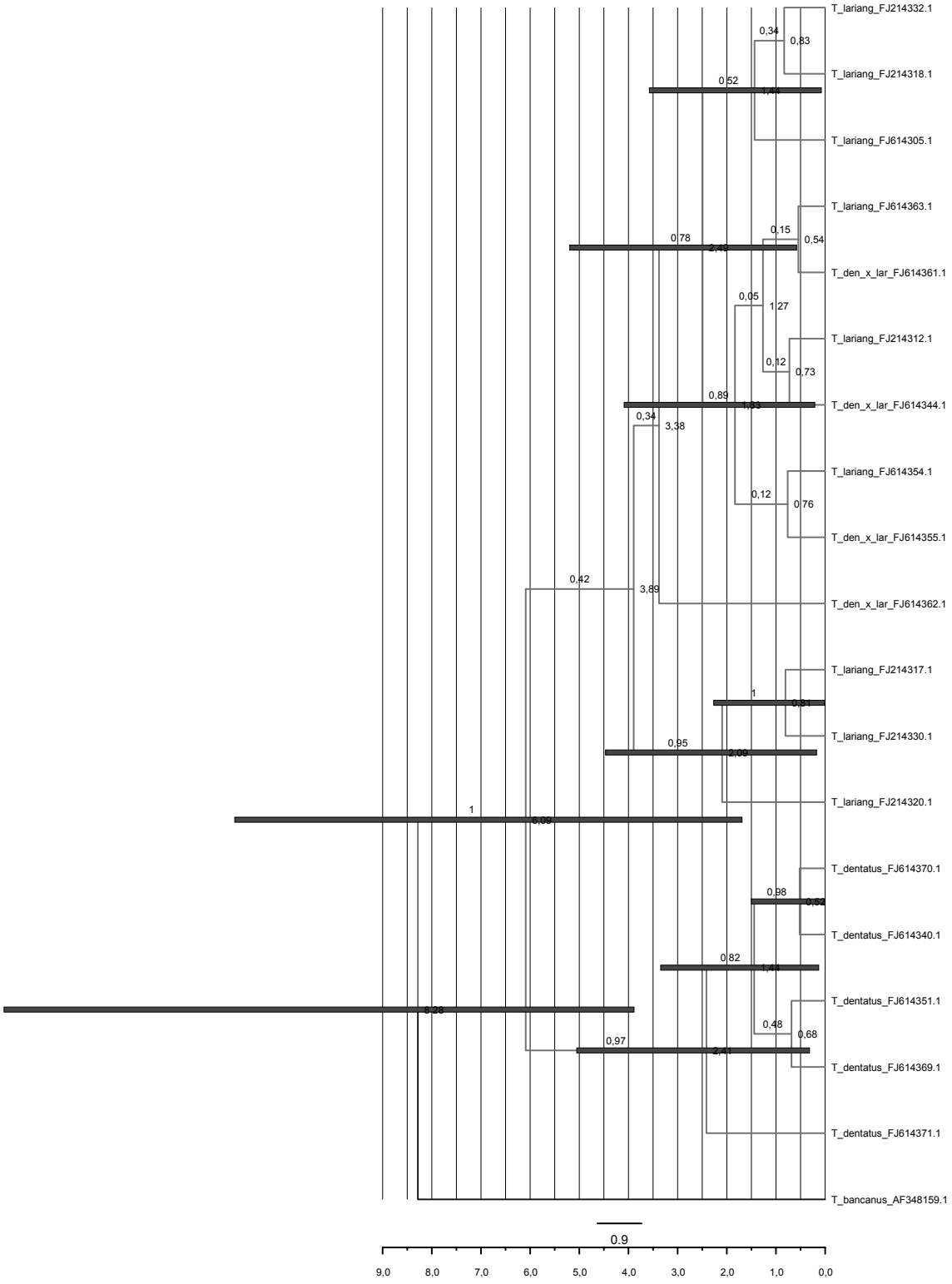
Dataset 14


L_macrocephalus_AY313717.1
L_macrocephalus_AY313716.1
L_macrocephalus_AY313713.1
L_macrocephalus_AY313714.1
L_macrocephalus_AY313715.1
L_macrocephalus_AY313718.1
L_visayanus_AY313720.1
L_visayanus_AY313719.1
L_woodworthi_AY313710.1
L_woodworthi_AY313709.1
L_woodworthi_AY313711.1
L_woodworthi_AY313708.1
L_woodworthi_AY313707.1
L_woodworthi_AY313712.1
L_modestus_AY313763.1
L_modestus_AY313765.1
L_modestus_AY313766.1
L_modestus_AY313764.1
L_modestus_AY313735.1
L_modestus_AY313736.1
L_cf_magnus_AY313704.1
L_magnus_AY313705.1
L_magnus_AY313706.1
L_magnus_AY313703.1
L_sp1_AY313762.1
L_sp1_AY313761.1
L_spV_AY313750.1
L_spV_AY313752.1
L_spV_AY313755.1
L_spV_AY313753.1
L_spV_AY313754.1
L_spV_AY313756.1
L_spV_AY313737.1
L_spV_AY313738.1
L_heinrichi_AY313748.1
L_heinrichi_AY313749.1
L_heinrichi_AY313747.1
L_spT_AY313759.1
L_spT_AY313757.1
L_spT_AY313758.1
L_spT_AY313760.1
L_spG_AY313734.1
L_spG_AY313733.1
L_spG_AY313732.1
L_spG_AY313731.1
L_microtympanum_AY313743.1
L_microtympanum_AY313745.1
L_microtympanum_AY313746.1
L_spD_AY313740.1
L_spD_AY313739.1
L_sp2_AY313741.1
L_sp2_AY313742.1
L_arathooni_AY313744.1
L_acanthi_AY313723.1
L_acanthi_AY313722.1
L_acanthi_AY313724.1
L_acanthi_AY313721.1
L_leytensis_AY313702.1
L_leytensis_AY313700.1
L_leytensis_AY313699.1
L_leytensis_AY313701.1
L_leytensis_AY313695.1
L_leytensis_AY313696.1
L_leytensis_AY313697.1
L_leytensis_AY313698.1
L_spI_AY313727.1
L_spI_AY313726.1
L_spI_AY313725.1
L_spI_AY313730.1
L_spI_AY313729.1
L_spI_AY313728.1
L_leporinus_AY313691.1
L_paramacrodon_AY313690.1
L_paramacrodon_AY313689.1
L_malesianus_AY313692.1
L_parvus_AY313694.1
L_kadarsani_AY313693.1
L_microdiscus_AY313688.1
L_kuhlii_AY313687.1
L_kuhlii_AY313686.1
F_vittigera_AY313683.1
H_rugulosus_AY313685.1
O_laevis_AY313684.1
3.0
20,0
15,0
10,0
5,0
0,0

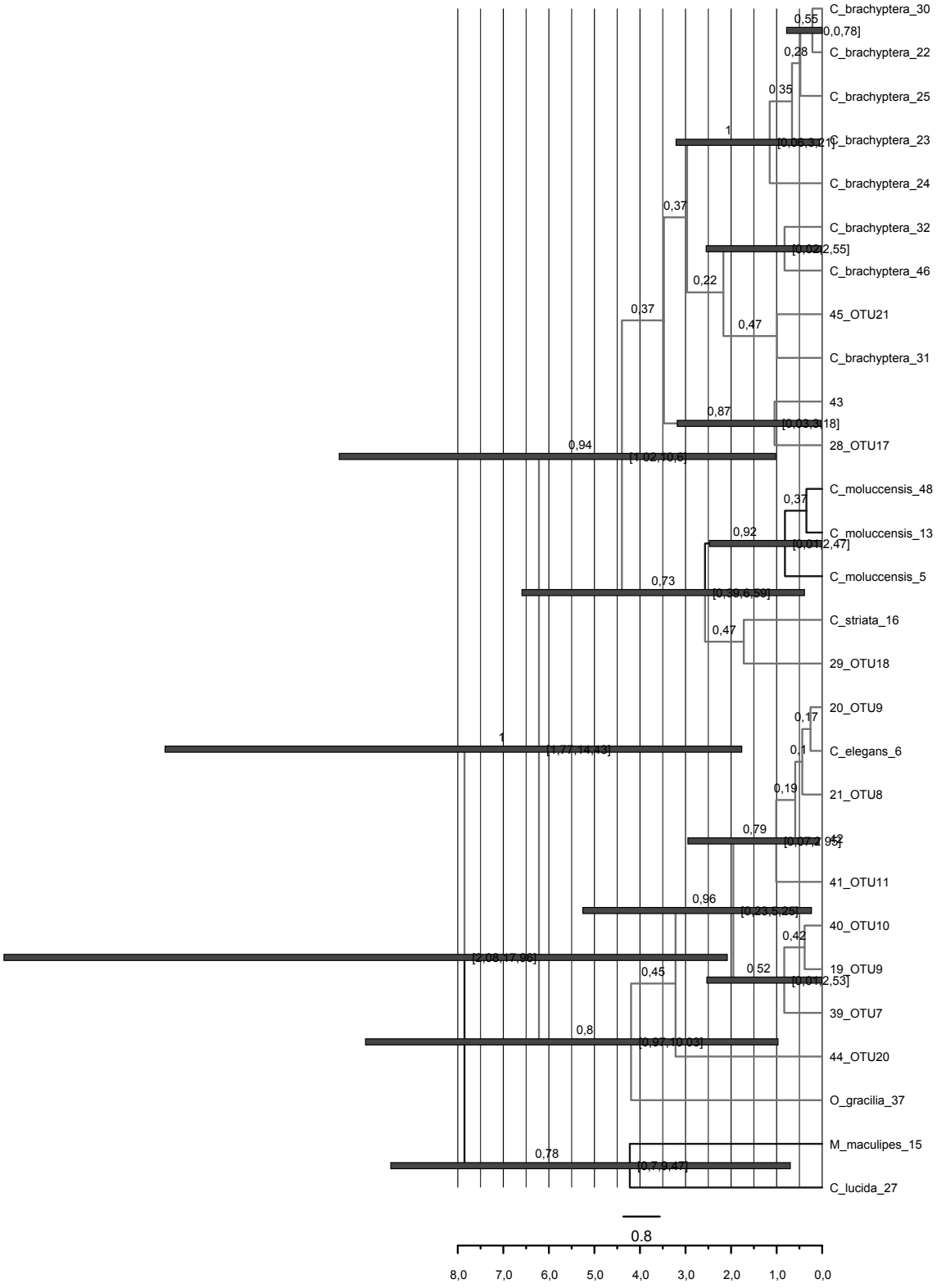
Dataset 15


L_macrocephalus_AY313718.1
L_macrocephalus_AY313715.1
L_macrocephalus_AY313713.1
L_macrocephalus_AY313714.1
L_macrocephalus_AY313716.1
L_macrocephalus_AY313717.1
L_visayanus_AY313719.1
L_visayanus_AY313720.1
L_magnus_AY313705.1
L_magnus_AY313706.1
L_cf_magnus_AY313704.1
L_woodworthi_AY313709.1
L_woodworthi_AY313711.1
L_woodworthi_AY313710.1
L_woodworthi_AY313712.1
L_woodworthi_AY313708.1
L_woodworthi_AY313707.1
L_modestus_AY313735.1
L_modestus_AY313736.1
L_modestus_AY313765.1
L_modestus_AY313763.1
L_modestus_AY313766.1
L_modestus_AY313764.1
L_magnus_AY313703.1
L_sp1_AY313762.1
L_sp1_AY313761.1
L_spT_AY313759.1
L_spT_AY313760.1
L_spT_AY313757.1
L_spT_AY313758.1
L_spV_AY313750.1
L_spV_AY313751.1
L_spV_AY313752.1
L_spV_AY313737.1
L_spV_AY313738.1
L_spV_AY313754.1
L_spV_AY313753.1
L_spV_AY313756.1
L_spV_AY313755.1
L_heinrichi_AY313748.1
L_heinrichi_AY313749.1
L_heinrichi_AY313747.1
L_spG_AY313732.1
L_spG_AY313733.1
L_spG_AY313731.1
L_microtympanum_AY313745.1
L_microtympanum_AY313743.1
L_microtympanum_AY313746.1
L_spD_AY313740.1
L_spD_AY313739.1
L_sp2_AY313741.1
L_sp2_AY313742.1
L_leytensis_AY313702.1
L_leytensis_AY313699.1
L_leytensis_AY313701.1
L_leytensis_AY313698.1
L_leytensis_AY313695.1
L_leytensis_AY313696.1
L_leytensis_AY313697.1
L_arathooni_AY313744.1
L_acanthi_AY313723.1
L_acanthi_AY313724.1
L_acanthi_AY313722.1
L_acanthi_AY313721.1
L_spI_AY313726.1
L_spI_AY313727.1
L_spI_AY313725.1
L_spI_AY313730.1
L_spI_AY313729.1
L_spI_AY313728.1
L_parvus_AY313694.1
L_paramacrodon_AY313690.1
L_paramacrodon_AY313689.1
L_leporinus_AY313691.1
L_kadarsani_AY313693.1
L_microdiscus_AY313688.1
L_kuhlii_AY313687.1
L_malesianus_AY313692.1
L_kuhlii_AY313686.1
O_laevis_AY313684.1
H_rugulosus_AY313685.1
F_vittigera_AY313683.1
30
20,0
15,0
10,0
5,0
0,0

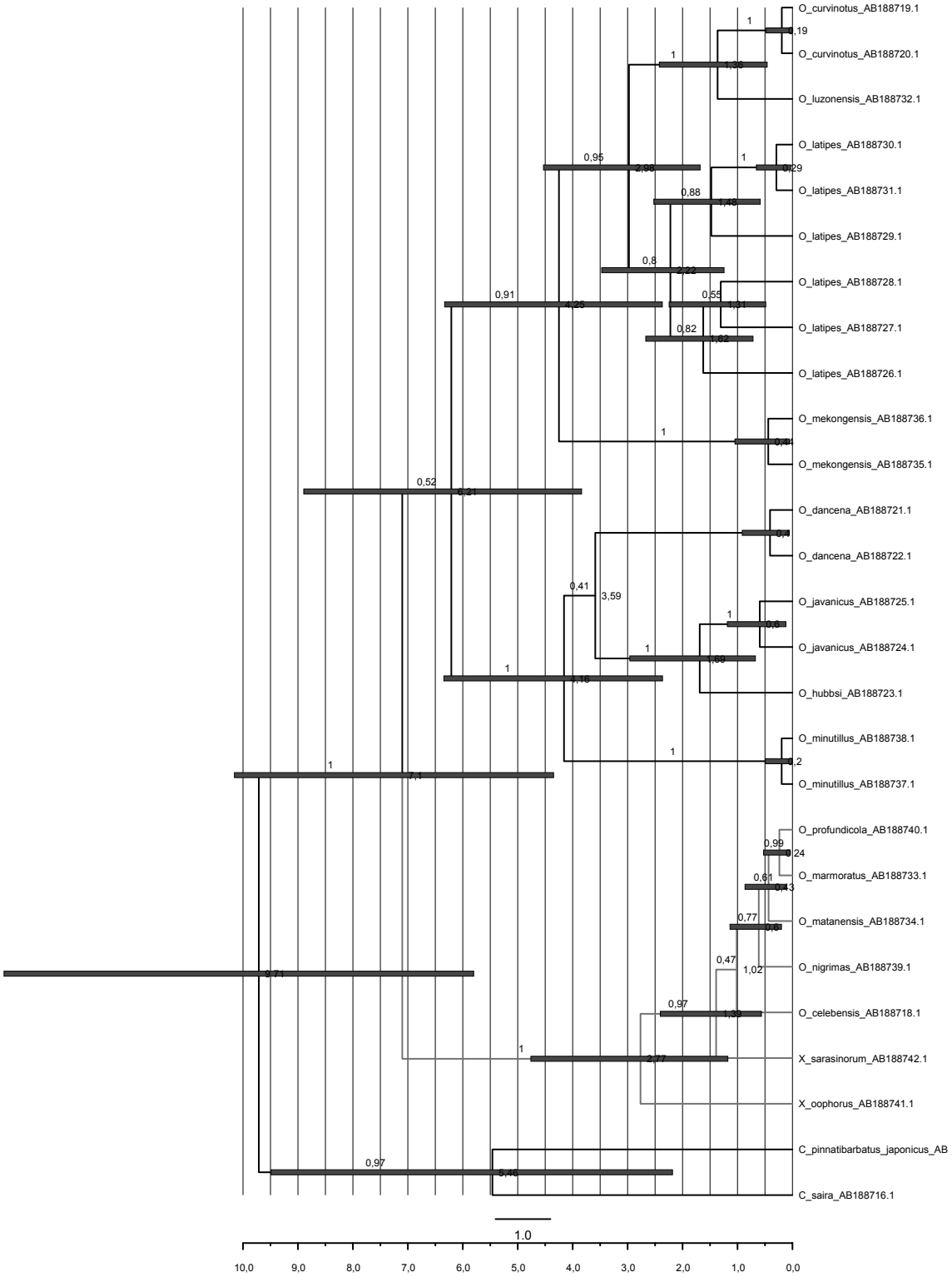
Dataset 16


T_lariang_FJ214332.1
0,34
0,83
T_lariang_FJ214318.1
0,52
T_lariang_FJ614305.1
T_lariang_FJ614363.1
0,15
0,78
0,54
T_den_x_lar_FJ614361.1
0,05
1,27
T_lariang_FJ214312.1
0,12
0,73
0,89
T_den_x_lar_FJ614344.1
0,34
3,38
T_lariang_FJ614354.1
0,12
0,76
T_den_x_lar_FJ614355.1
0,42
3,89
T_den_x_lar_FJ614362.1
T_lariang_FJ214317.1
1
T_lariang_FJ214330.1
0,95
T_lariang_FJ214320.1
1
T_dentatus_FJ614370.1
0,98
T_dentatus_FJ614340.1
0,82
T_dentatus_FJ614351.1
0,48
0,68
0,97
T_dentatus_FJ614369.1
T_dentatus_FJ614371.1
T_bancanus_AF348159.1
0.9
9,0
8,0
7,0
6,0
5,0
4,0
3,0
2,0
1,0
0,0

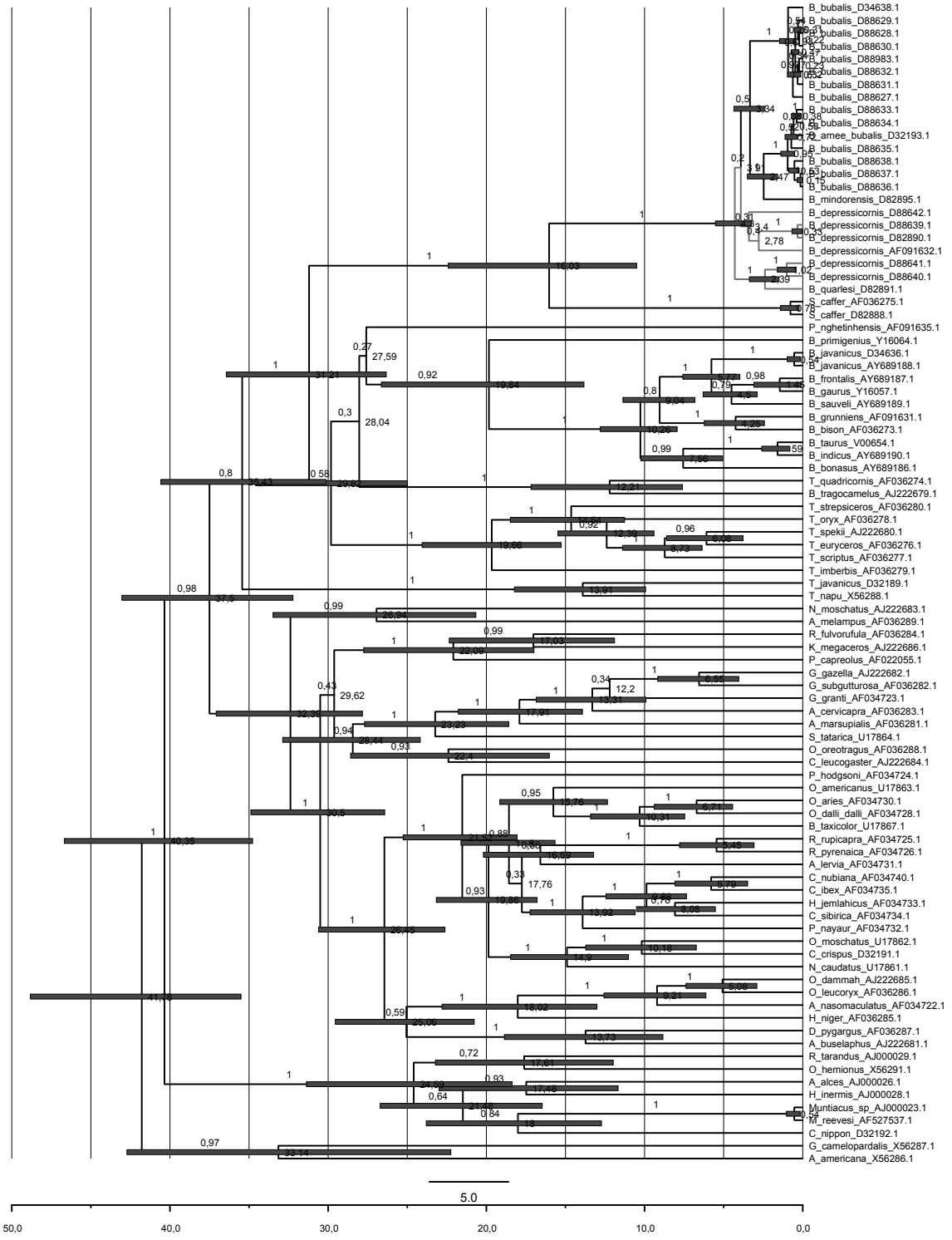
Dataset 17


C_brachyptera_30
0,55
0,0,78]
0,28
C_brachyptera_22
0,35
C_brachyptera_25
1
C_brachyptera_23
C_brachyptera_24
0,37
C_brachyptera_32
C_brachyptera_46
0,22
0,37
45_OTU21
0,47
C_brachyptera_31
43
0,87
28_OTU17
0,94
C_moluccensis_48
0,37
C_moluccensis_13
0,92
C_moluccensis_5
0,73
C_striata_16
0,47
29_OTU18
20_OTU9
0,17
1
C_elegans_6
0,1
0,19
21_OTU8
0,79
41_OTU11
0,96
40_OTU10
0,42
19_OTU9
0,52
0,45
39_OTU7
0,8
44_OTU20
O_gracilia_37
M_maculipes_15
0,78
C_lucida_27
0.8
8,0
7,0
6,0
5,0
4,0
3,0
2,0
1,0
0,0

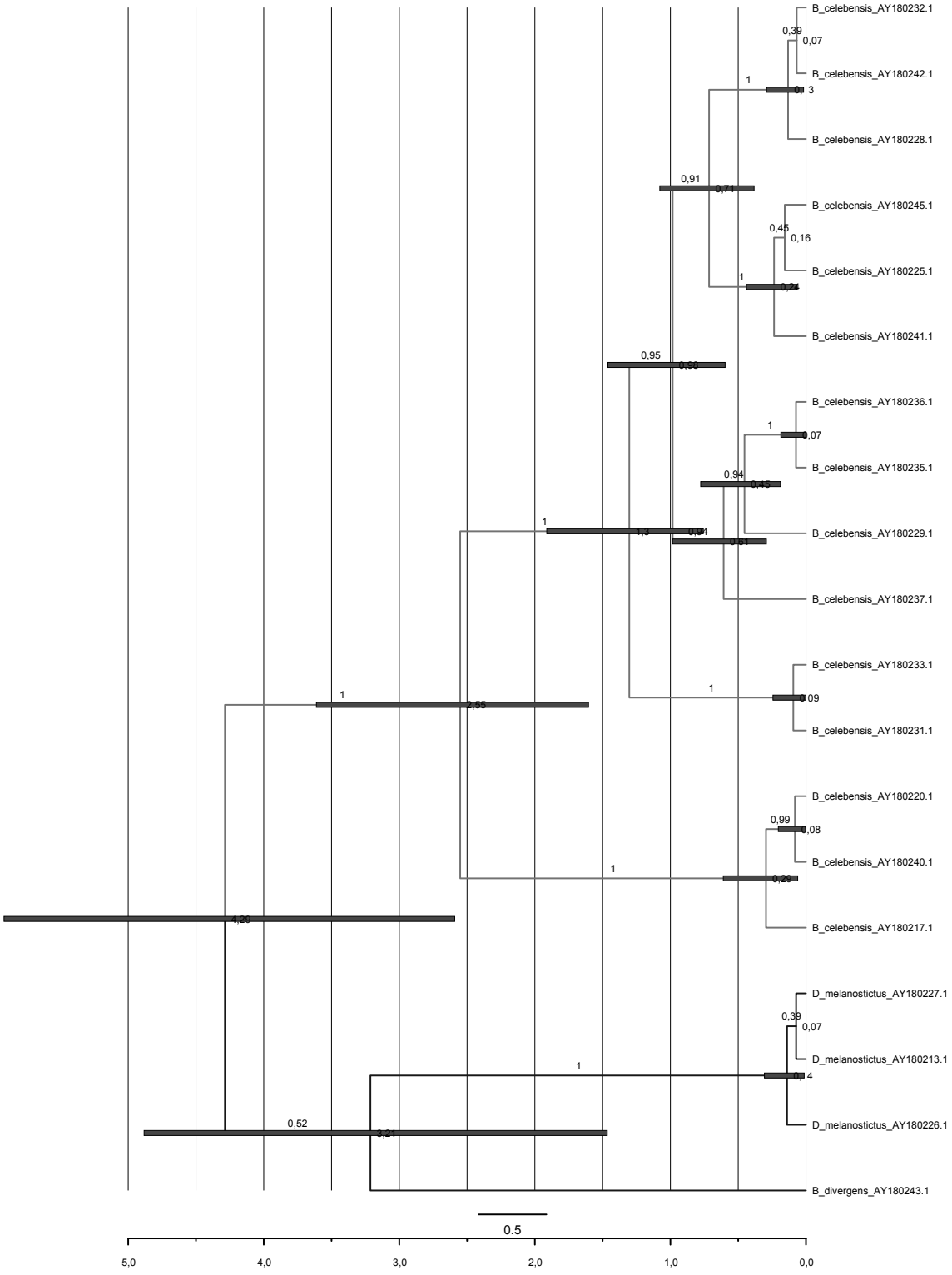
Dataset 18


O_curvinotus_AB188719.1
O_curvinotus_AB188720.1
O_luzonensis_AB188732.1
O_latipes_AB188730.1
O_latipes_AB188731.1
O_latipes_AB188729.1
O_latipes_AB188728.1
O_latipes_AB188727.1
O_latipes_AB188726.1
O_mekongensis_AB188736.1
O_mekongensis_AB188735.1
O_dancena_AB188721.1
O_dancena_AB188722.1
O_javanicus_AB188725.1
O_javanicus_AB188724.1
O_hubbsi_AB188723.1
O_minutillus_AB188738.1
O_minutillus_AB188737.1
O_profundicola_AB188740.1
O_marmoratus_AB188733.1
O_matanensis_AB188734.1
O_nigrimas_AB188739.1
O_celebensis_AB188718.1
X_sarasinorum_AB188742.1
X_oophorus_AB188741.1
C_pinnatibarbatus_japonicus_AB
C_saira_AB188716.1
1.0
10,0
9,0
8,0
7,0
6,0
5,0
4,0
3,0
2,0
1,0
0,0

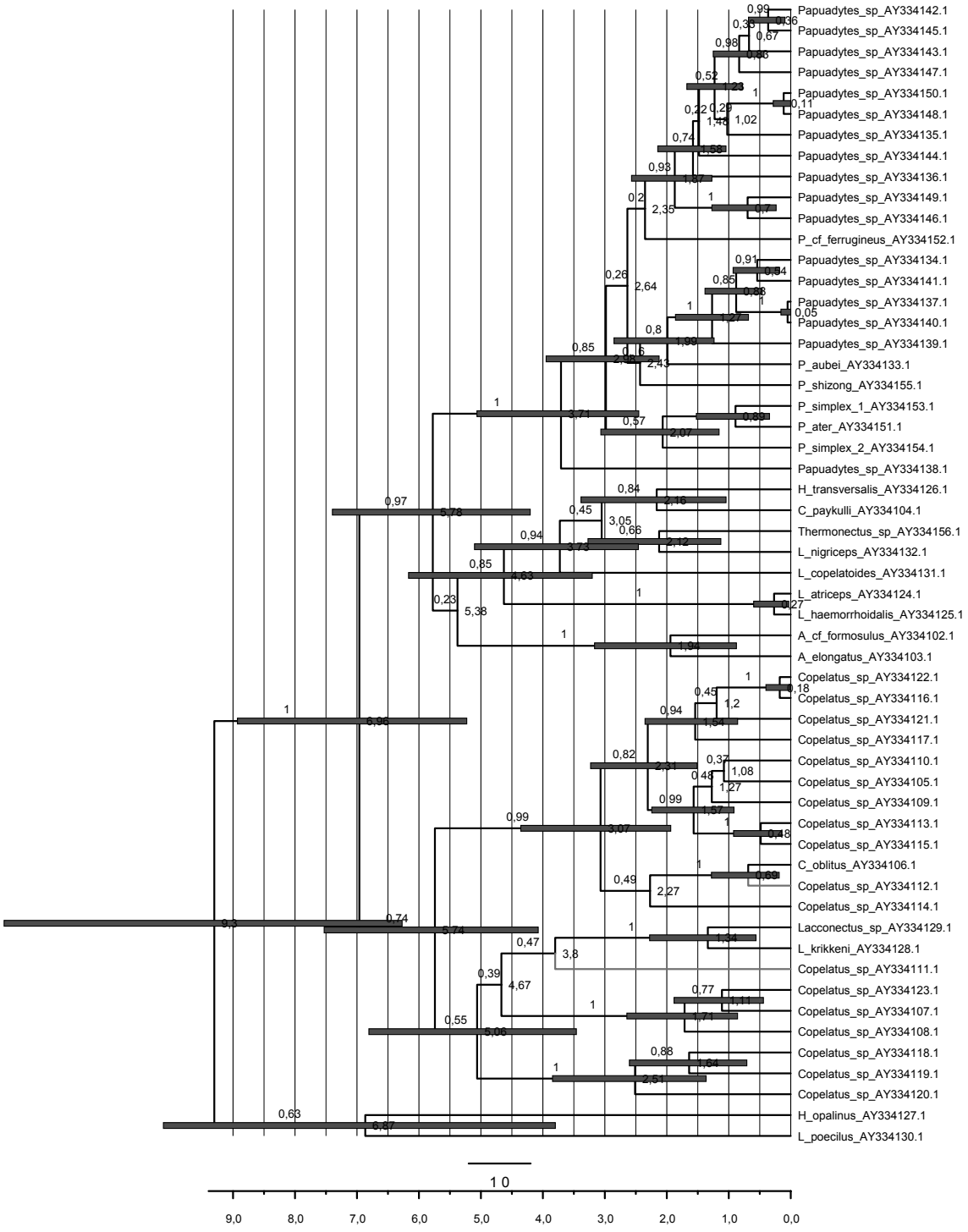
Dataset 19


B_bubalis_D34638.1
B_bubalis_D88629.1
B_bubalis_D88628.1
B_bubalis_D88630.1
B_bubalis_D88983.1
B_bubalis_D88632.1
B_bubalis_D88631.1
B_bubalis_D88627.1
B_bubalis_D88633.1
B_bubalis_D88634.1
B_arnee_bubalis_D32193.1
B_bubalis_D88635.1
B_bubalis_D88638.1
B_bubalis_D88637.1
B_bubalis_D88636.1
B_mindorensis_D82895.1
B_depressicornis_D88642.1
B_depressicornis_D88639.1
B_depressicornis_D82890.1
B_depressicornis_AF091632.1
B_depressicornis_D88641.1
B_depressicornis_D88640.1
B_quarlesi_D82891.1
S_caffer_AF036275.1
S_caffer_D82888.1
P_nghetinhensis_AF091635.1
B_primigenius_Y16064.1
B_javanicus_D34636.1
B_javanicus_AY689188.1
B_frontalis_AY689187.1
B_gaurus_Y16057.1
B_sauveli_AY689189.1
B_grunniens_AF091631.1
B_bison_AF036273.1
B_taurus_V00654.1
B_indicus_AY689190.1
B_bonasus_AY689186.1
T_quadricornis_AF036274.1
B_tragocamelus_AJ222679.1
T_strepsiceros_AF036280.1
T_oryx_AF036278.1
T_spekii_AJ222680.1
T_euryceros_AF036276.1
T_scriptus_AF036277.1
T_imberbis_AF036279.1
T_javanicus_D32189.1
T_napu_X56288.1
N_moschatus_AJ222683.1
A_melampus_AF036289.1
R_fulvorufula_AF036284.1
K_megaceros_AJ222686.1
P_capreolus_AF022055.1
G_gazella_AJ222682.1
G_subgutturosa_AF036282.1
G_granti_AF034723.1
A_cervicapra_AF036283.1
A_marsupialis_AF036281.1
S_tatarica_U17864.1
O_oreotragus_AF036288.1
C_leucogaster_AJ222684.1
P_hodgsoni_AF034724.1
O_americanus_U17863.1
O_aries_AF034730.1
O_dalli_dalli_AF034728.1
B_taxicolor_U17867.1
R_rupicapra_AF034725.1
R_pyrenaica_AF034726.1
A_lervia_AF034731.1
C_nubiana_AF034740.1
C_ibex_AF034735.1
H_jemlahicus_AF034733.1
C_sibirica_AF034734.1
P_nayaur_AF034732.1
O_moschatus_U17862.1
C_crispus_D32191.1
N_caudatus_U17861.1
O_dammah_AJ222685.1
O_leucoryx_AF036286.1
A_nasomaculatus_AF034722.1
H_niger_AF036285.1
D_pygargus_AF036287.1
A_buselaphus_AJ222681.1
R_tarandus_AJ000029.1
O_hemionus_X56291.1
A_alces_AJ000026.1
H_inermis_AJ000028.1
Muntiacus_sp_AJ000023.1
M_reevesi_AF527537.1
C_nippon_D32192.1
G_camelopardalis_X56287.1
A_americana_X56286.1
5.0
50,0
40,0
30,0
20,0
10,0
0,0

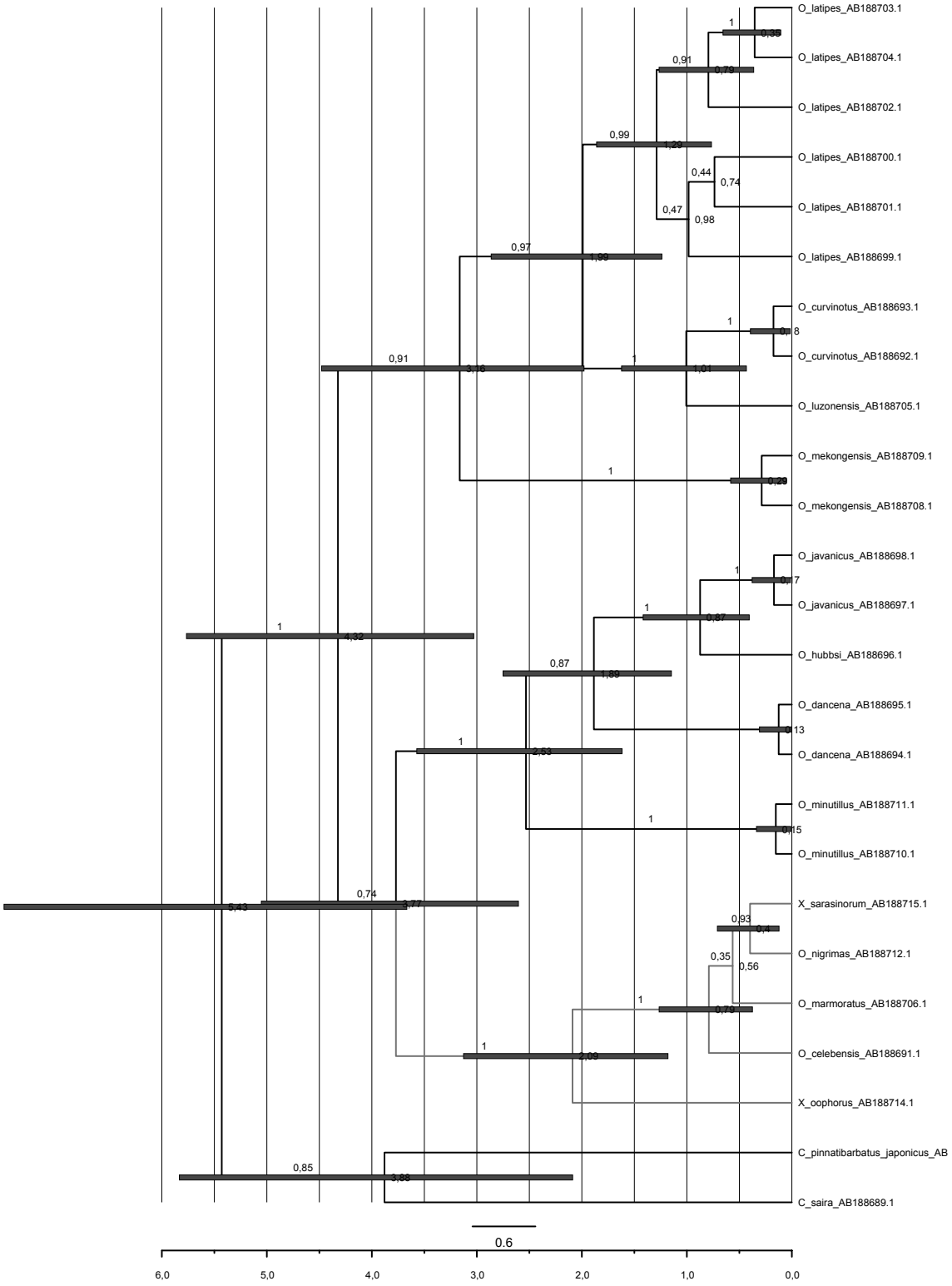
Dataset 20


B_celebensis_AY180232.1
0,39
0,07
B_celebensis_AY180242.1
1
B_celebensis_AY180228.1
0,91
B_celebensis_AY180245.1
0,45
0,16
B_celebensis_AY180225.1
1
B_celebensis_AY180241.1
0,95
B_celebensis_AY180236.1
1
0,07
B_celebensis_AY180235.1
0,94
B_celebensis_AY180229.1
1
B_celebensis_AY180237.1
B_celebensis_AY180233.1
1
1
B_celebensis_AY180231.1
B_celebensis_AY180220.1
0,99
B_celebensis_AY180240.1
1
B_celebensis_AY180217.1
D_melanostictus_AY180227.1
0,39
0,07
D_melanostictus_AY180213.1
1
0,52
D_melanostictus_AY180226.1
B_divergens_AY180243.1
0.5
5,0
4,0
3,0
2,0
1,0
0,0

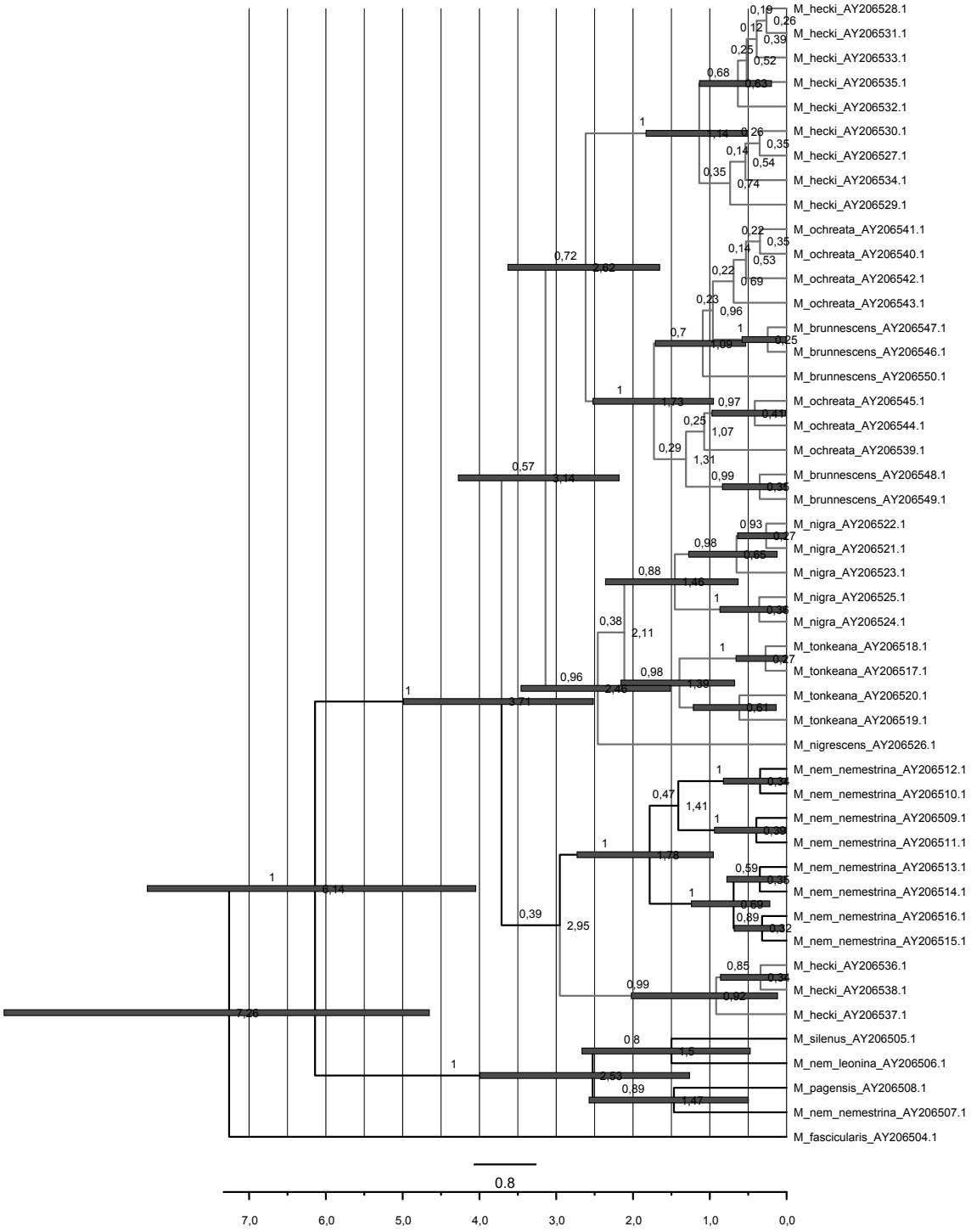
Dataset 21


Papuadytes_sp_AY334142.1
Papuadytes_sp_AY334145.1
Papuadytes_sp_AY334143.1
Papuadytes_sp_AY334147.1
Papuadytes_sp_AY334150.1
Papuadytes_sp_AY334148.1
Papuadytes_sp_AY334135.1
Papuadytes_sp_AY334144.1
Papuadytes_sp_AY334136.1
Papuadytes_sp_AY334149.1
Papuadytes_sp_AY334146.1
P_cf_ferrugineus_AY334152.1
Papuadytes_sp_AY334134.1
Papuadytes_sp_AY334141.1
Papuadytes_sp_AY334137.1
Papuadytes_sp_AY334140.1
Papuadytes_sp_AY334139.1
P_aubei_AY334133.1
P_shizong_AY334155.1
P_simplex_1_AY334153.1
P_ater_AY334151.1
P_simplex_2_AY334154.1
Papuadytes_sp_AY334138.1
H_transversalis_AY334126.1
C_paykulli_AY334104.1
Thermonectus_sp_AY334156.1
L_nigriceps_AY334132.1
L_copelatoides_AY334131.1
L_atriceps_AY334124.1
L_haemorrhoidalis_AY334125.1
A_cf_formosulus_AY334102.1
A_elongatus_AY334103.1
Copelatus_sp_AY334122.1
Copelatus_sp_AY334116.1
Copelatus_sp_AY334121.1
Copelatus_sp_AY334117.1
Copelatus_sp_AY334110.1
Copelatus_sp_AY334105.1
Copelatus_sp_AY334109.1
Copelatus_sp_AY334113.1
Copelatus_sp_AY334115.1
C_oblitus_AY334106.1
Copelatus_sp_AY334112.1
Copelatus_sp_AY334114.1
Lacconectus_sp_AY334129.1
L_krikkeni_AY334128.1
Copelatus_sp_AY334111.1
Copelatus_sp_AY334123.1
Copelatus_sp_AY334107.1
Copelatus_sp_AY334108.1
Copelatus_sp_AY334118.1
Copelatus_sp_AY334119.1
Copelatus_sp_AY334120.1
H_opalinus_AY334127.1
L_poecilus_AY334130.1
1 0
9,0
8,0
7,0
6,0
5,0
4,0
3,0
2,0
1,0
0,0

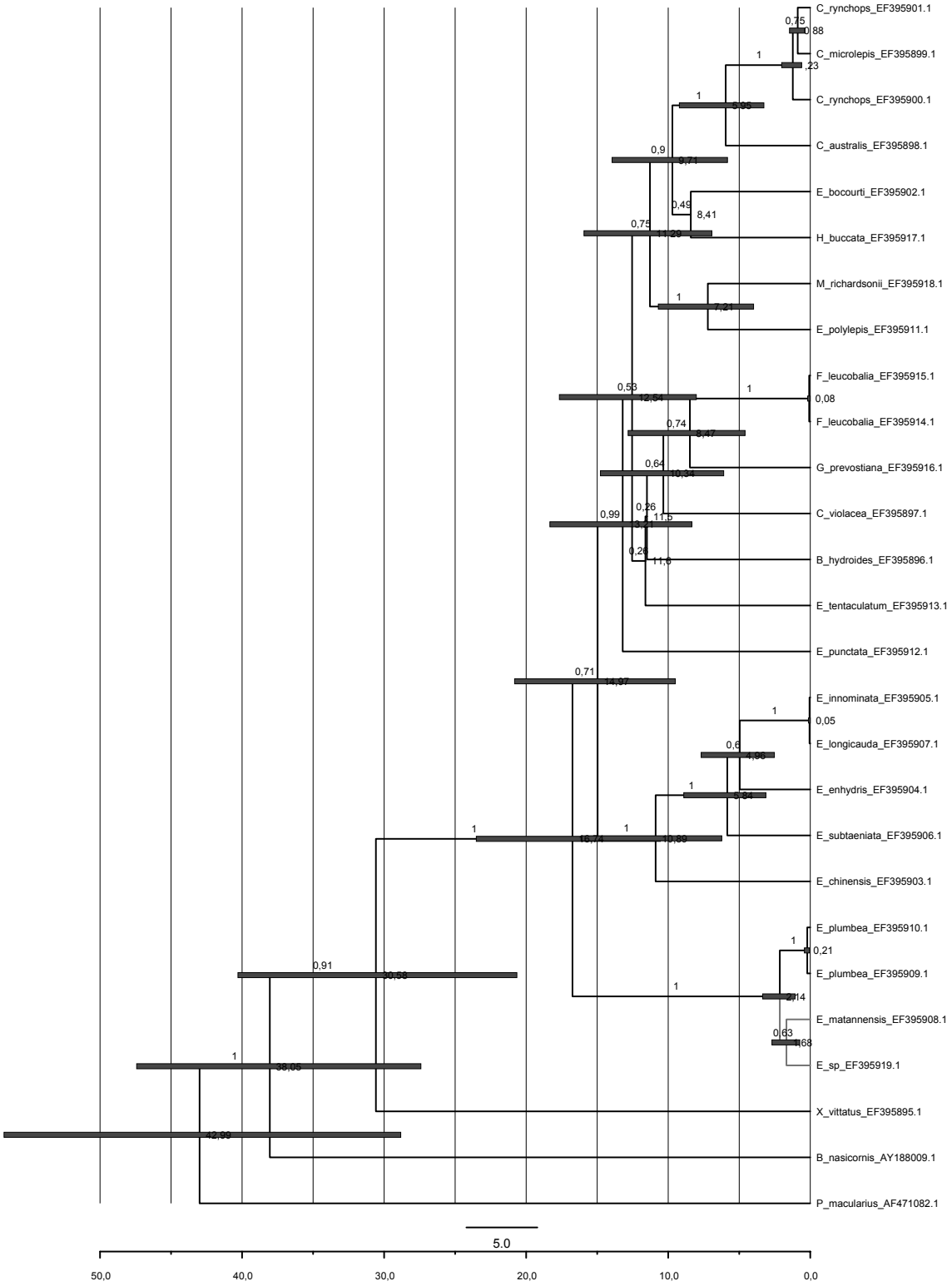
Dataset 22


O_latipes_AB188703.1
O_latipes_AB188704.1
O_latipes_AB188702.1
O_latipes_AB188700.1
O_latipes_AB188701.1
O_latipes_AB188699.1
O_curvinotus_AB188693.1
O_curvinotus_AB188692.1
O_luzonensis_AB188705.1
O_mekongensis_AB188709.1
O_mekongensis_AB188708.1
O_javanicus_AB188698.1
O_javanicus_AB188697.1
O_hubbsi_AB188696.1
O_dancena_AB188695.1
O_dancena_AB188694.1
O_minutillus_AB188711.1
O_minutillus_AB188710.1
X_sarasinorum_AB188715.1
O_nigrimas_AB188712.1
O_marmoratus_AB188706.1
O_celebensis_AB188691.1
X_oophorus_AB188714.1
C_pinnatibarbatus_japonicus_AB
C_saira_AB188689.1
0.6
6,0
5,0
4,0
3,0
2,0
1,0
0,0

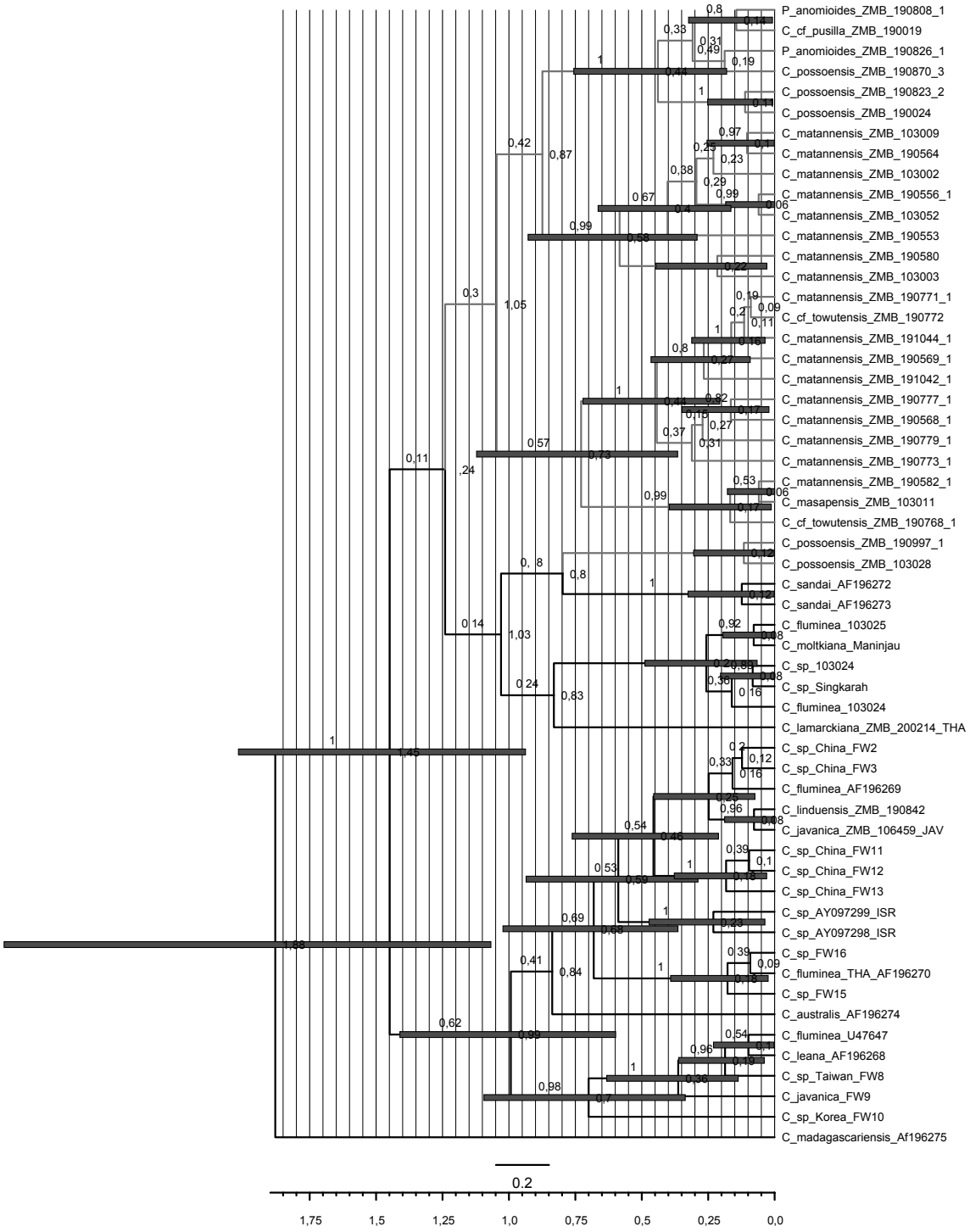
Dataset 23


M_hecki_AY206528.1
M_hecki_AY206531.1
M_hecki_AY206533.1
M_hecki_AY206535.1
M_hecki_AY206532.1
M_hecki_AY206530.1
M_hecki_AY206527.1
M_hecki_AY206534.1
M_hecki_AY206529.1
M_ochreata_AY206541.1
M_ochreata_AY206540.1
M_ochreata_AY206542.1
M_ochreata_AY206543.1
M_brunnescens_AY206547.1
M_brunnescens_AY206546.1
M_brunnescens_AY206550.1
M_ochreata_AY206545.1
M_ochreata_AY206544.1
M_ochreata_AY206539.1
M_brunnescens_AY206548.1
M_brunnescens_AY206549.1
M_nigra_AY206522.1
M_nigra_AY206521.1
M_nigra_AY206523.1
M_nigra_AY206525.1
M_nigra_AY206524.1
M_tonkeana_AY206518.1
M_tonkeana_AY206517.1
M_tonkeana_AY206520.1
M_tonkeana_AY206519.1
M_nigrescens_AY206526.1
M_nem_nemestrina_AY206512.1
M_nem_nemestrina_AY206510.1
M_nem_nemestrina_AY206509.1
M_nem_nemestrina_AY206511.1
M_nem_nemestrina_AY206513.1
M_nem_nemestrina_AY206514.1
M_nem_nemestrina_AY206516.1
M_nem_nemestrina_AY206515.1
M_hecki_AY206536.1
M_hecki_AY206538.1
M_hecki_AY206537.1
M_silenus_AY206505.1
M_nem_leonina_AY206506.1
M_pagensis_AY206508.1
M_nem_nemestrina_AY206507.1
M_fascicularis_AY206504.1
0.8
7,0
6,0
5,0
4,0
3,0
2,0
1,0
0,0

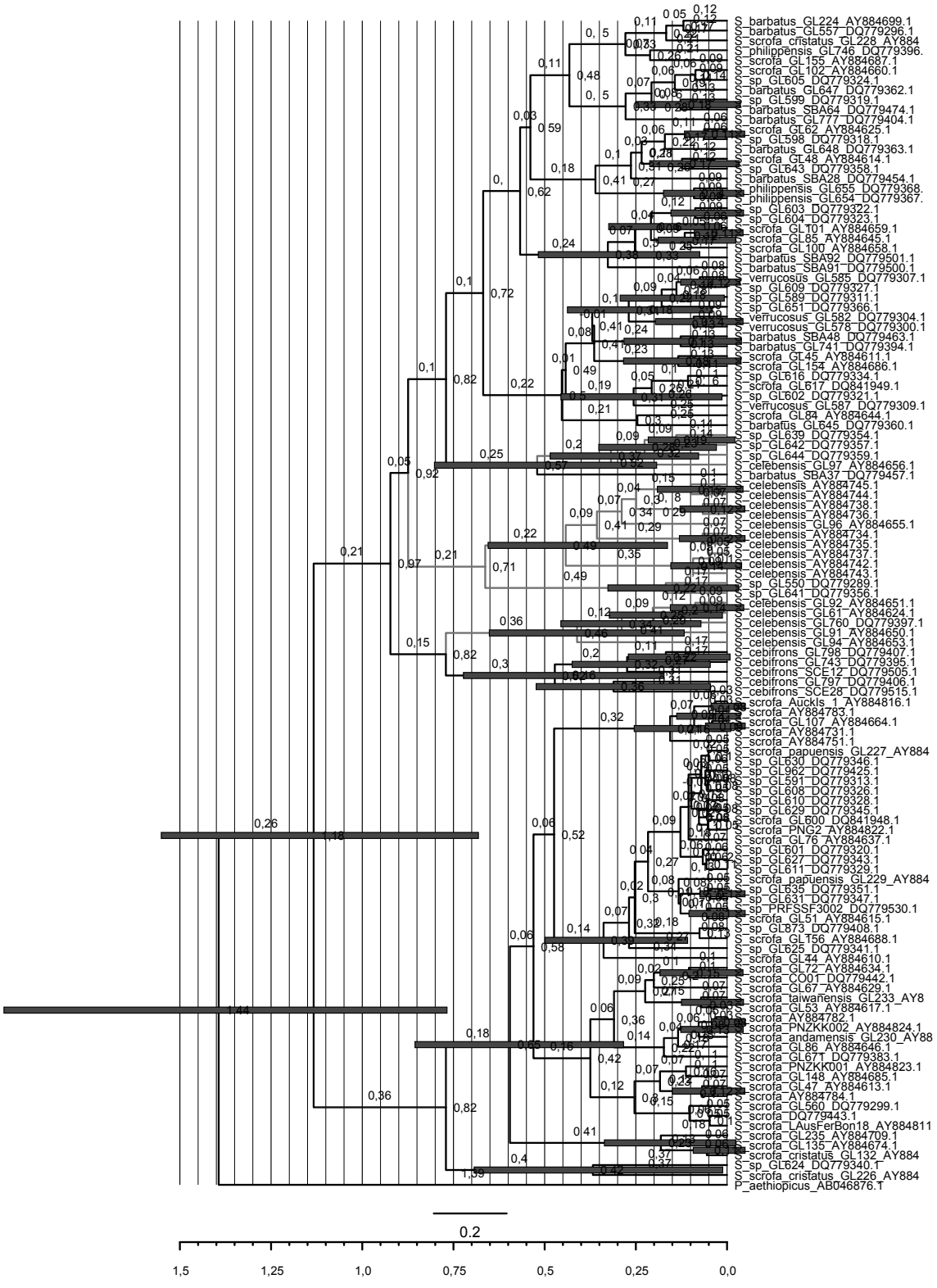
Dataset 24


C_rynchops_EF395901.1
C_microlepis_EF395899.1
C_rynchops_EF395900.1
C_australis_EF395898.1
E_bocourti_EF395902.1
H_buccata_EF395917.1
M_richardsonii_EF395918.1
E_polylepis_EF395911.1
F_leucobalia_EF395915.1
F_leucobalia_EF395914.1
G_prevostiana_EF395916.1
C_violacea_EF395897.1
B_hydroides_EF395896.1
E_tentaculatum_EF395913.1
E_punctata_EF395912.1
E_innominata_EF395905.1
E_longicauda_EF395907.1
E_enhydris_EF395904.1
E_subtaeniata_EF395906.1
E_chinensis_EF395903.1
E_plumbea_EF395910.1
E_plumbea_EF395909.1
E_matannensis_EF395908.1
E_sp_EF395919.1
X_vittatus_EF395895.1
B_nasicornis_AY188009.1
P_macularius_AF471082.1
5.0
50,0
40,0
30,0
20,0
10,0
0,0

Dataset 25


P_anomioides_ZMB_190808_1
C_cf_pusilla_ZMB_190019
P_anomioides_ZMB_190826_1
C_possoensis_ZMB_190870_3
C_possoensis_ZMB_190823_2
C_possoensis_ZMB_190024
C_matannensis_ZMB_103009
C_matannensis_ZMB_190564
C_matannensis_ZMB_103002
C_matannensis_ZMB_190556_1
C_matannensis_ZMB_103052
C_matannensis_ZMB_190553
C_matannensis_ZMB_190580
C_matannensis_ZMB_103003
C_matannensis_ZMB_190771_1
C_cf_towutensis_ZMB_190772
C_matannensis_ZMB_191044_1
C_matannensis_ZMB_190569_1
C_matannensis_ZMB_191042_1
C_matannensis_ZMB_190777_1
C_matannensis_ZMB_190568_1
C_matannensis_ZMB_190779_1
C_matannensis_ZMB_190773_1
C_matannensis_ZMB_190582_1
C_masapensis_ZMB_103011
C_cf_towutensis_ZMB_190768_1
C_possoensis_ZMB_190997_1
C_possoensis_ZMB_103028
C_sandai_AF196272
C_sandai_AF196273
C_fluminea_103025
C_moltkiana_Maninjau
C_sp_103024
C_sp_Singkarah
C_fluminea_103024
C_lamarckiana_ZMB_200214_THA
C_sp_China_FW2
C_sp_China_FW3
C_fluminea_AF196269
C_linduensis_ZMB_190842
C_javanica_ZMB_106459_JAV
C_sp_China_FW11
C_sp_China_FW12
C_sp_China_FW13
C_sp_AY097299_ISR
C_sp_AY097298_ISR
C_sp_FW16
C_fluminea_THA_AF196270
C_sp_FW15
C_australis_AF196274
C_fluminea_U47647
C_leana_AF196268
C_sp_Taiwan_FW8
C_javanica_FW9
C_sp_Korea_FW10
C_madagascariensis_Af196275
0.2
1,75
1,5
1,25
1,0
0,75
0,5
0,25
0,0

Dataset 26


S_barbatus_GL224_AY884699.1
S_barbatus_GL557_DQ779296.1
S_scrofa_cristatus_GL228_AY884
S_philippensis_GL746_DQ779396.
S_scrofa_GL155_AY884687.1
S_scrofa_GL102_AY884660.1
S_sp_GL605_DQ779324.1
S_barbatus_GL647_DQ779362.1
S_sp_GL599_DQ779319.1
S_barbatus_SBA64_DQ779474.1
S_barbatus_GL777_DQ779404.1
S_scrofa_GL62_AY884625.1
S_sp_GL598_DQ779318.1
S_barbatus_GL648_DQ779363.1
S_scrofa_GL48_AY884614.1
S_sp_GL643_DQ779358.1
S_barbatus_SBA28_DQ779454.1
S_philippensis_GL655_DQ779368.
S_philippensis_GL654_DQ779367.
S_sp_GL603_DQ779322.1
S_sp_GL604_DQ779323.1
S_scrofa_GL101_AY884659.1
S_scrofa_GL85_AY884645.1
S_scrofa_GL100_AY884658.1
S_barbatus_SBA92_DQ779501.1
S_barbatus_SBA91_DQ779500.1
S_verrucosus_GL585_DQ779307.1
S_sp_GL609_DQ779327.1
S_sp_GL589_DQ779311.1
S_sp_GL651_DQ779366.1
S_verrucosus_GL582_DQ779304.1
S_verrucosus_GL578_DQ779300.1
S_barbatus_SBA48_DQ779463.1
S_barbatus_GL741_DQ779394.1
S_scrofa_GL45_AY884611.1
S_scrofa_GL154_AY884686.1
S_sp_GL616_DQ779334.1
S_scrofa_GL617_DQ841949.1
S_sp_GL602_DQ779321.1
S_verrucosus_GL587_DQ779309.1
S_scrofa_GL8A_AY884644.1
S_barbatus_GL645_DQ779360.1
S_sp_GL639_DQ779354.1
S_sp_GL642_DQ779357.1
S_sp_GL644_DQ779359.1
S_celebensis_GL97_AY884656.1
S_barbatus_SBA37_DQ779457.1
S_celebensis_AY884745.1
S_celebensis_AY884744.1
S_celebensis_AY884738.1
S_celebensis_AY884736.1
S_celebensis_GL96_AY884655.1
S_celebensis_AY884734.1
S_celebensis_AY884735.1
S_celebensis_AY884737.1
S_celebensis_AY884742.1
S_celebensis_AY884743.1
S_sp_GL550_DQ779289.1
S_sp_GL641_DQ779356.1
S_celebensis_GL92_AY884651.1
S_celebensis_GL61_AY884624.1
S_celebensis_GL760_DQ779397.1
S_celebensis_GL91_AY884650.1
S_celebensis_GL94_AY884653.1
S_cebifrons_GL798_DQ779407.1
S_cebifrons_GL743_DQ779395.1
S_cebifrons_SCE12_DQ779505.1
S_cebifrons_GL797_DQ779406.1
S_cebifrons_SCE28_DQ779515.1
S_scrofa_Auckls_1_AY884816.1
S_scrofa_AY884783.1
S_scrofa_GL107_AY884664.1
S_scrofa_AY884731.1
S_scrofa_AY884751.1
S_scrofa_papuensis_GL227_AY884
S_sp_GL630_DQ779346.1
S_sp_GL962_DQ779425.1
S_sp_GL591_DQ779313.1
S_sp_GL608_DQ779326.1
S_sp_GL610_DQ779328.1
S_sp_GL629_DQ779345.1
S_scrofa_GL600_DQ841948.1
S_scrofa_PNG2_AY884822.1
S_scrofa_GL76_AY884637.1
S_sp_GL601_DQ779320.1
S_sp_GL627_DQ779343.1
S_sp_GL611_DQ779329.1
S_scrofa_papuensis_GL229_AY884
S_sp_GL635_DQ779351.1
S_sp_GL631_DQ779347.1
S_sp_PRFSSF3002_DQ779530.1
S_scrofa_GL51_AY884615.1
S_sp_GL873_DQ779408.1
S_scrofa_GL156_AY884688.1
S_sp_GL625_DQ779341.1
S_scrofa_GL44_AY884610.1
S_scrofa_GL72_AY884634.1
S_scrofa_C001_DQ779442.1
S_scrofa_GL67_AY884629.1
S_scrofa_taiwanensis_GL233_AY8
S_scrofa_GL53_AY884617.1
S_scrofa_AY884782.1
S_scrofa_PNZKK002_AY884824.1
S_scrofa_andamensis_GL230_AY88
S_scrofa_GL86_AY884646.1
S_scrofa_GL671_DQ779383.1
S_scrofa_PNZKK001_AY884823.1
S_scrofa_GL148_AY884685.1
S_scrofa_GL47_AY884613.1
S_scrofa_AY884784.1
S_scrofa_GL560_DQ779299.1
S_scrofa_DQ779443.1
S_scrofa_LAusFerBon18_AY884811
S_scrofa_GL235_AY884709.1
S_scrofa_GL135_AY884674.1
S_scrofa_cristatus_GL132_AY884
S_sp_GL624_DQ779340.1
S_scrofa_cristatus_GL226_AY884
P_aethiopicus_AB046876.1
0.2
1,5
1,25
1,0
0,75
0,5
0,25
0,0

Dataset 27

Supplementary Information – Chapter 3

Table S1. List of studied specimens including distribution information and accession numbers.

Family	Species	Voucher no. or GenBank no.[1]	Locality/distribution (AS = Aquarium stock)	Accession no. ND2	Accession no. 12S-16S
Notocheiridae	*Iso rhotophilus*	ZSM (RS10)*	Australia	KJ667882	KJ667787
Atherinidae	*Atherinomorus endrachtensis*	ZSM (RS3)*	Indo-Pacific	KJ667867	KJ667772
	Atherinomorus lacunosus	ZSM (RS5)*	Japan	KJ667868	KJ667773
	Craterocephalus honoriae	ZSM (RS6)*	Australia	KJ667877	KJ667782
	Hypoatherina tsurugae	AP004420	Gregory River, Queensland, Australia	(Miya et al., 2003)	(Miya et al., 2003)
	Quirichthys sp.	ZSM 34172-1	Gregory River, Queensland, Australia	KJ667922	KJ667824
	Quirichthys sp.	ZSM 34172-3	Gregory River, Queensland, Australia	KJ667923	KJ667825
Atherinopsidae	*Atherinella panamensis*	ZSM (RS4)*	Panama (Eastern Pacific)	KJ667866	KJ667771
Bedotiidae	*Bedotia* sp.	ZSM 34169	Ankavia, Madagascar (AS)	KJ667870	KJ667775
	Bedotia sp.	ZSM 34168	Lazana, Madagascar (AS)	KJ667869	KJ667774
Melanotaeniidae	*Chilatherina bleheri*	ZSM 34140	Papua New Guinea (AS)	KJ667872	KJ667777
	Chilatherina campsi	ZSM 34177	Papua New Guinea (AS)	KJ667873	KJ667778
	Chilatherina fasciata	ZSM 34137	Clearwater Creek, Papua New Guinea (AS)	KJ667874	KJ667779
	Chilatherina fasciata	ZSM 34142	Lake Wanam, Papua New Guinea (AS)	KJ667876	KJ667781
	Chilatherina fasciata	ZSM 34141	Lake Sentani, West Papua, Indonesia (AS)	KJ667875	KJ667780
	Chilatherina sp.	ZSM 34143	Mamberamo River, West Papua, Indonesia (AS)	KJ667871	KJ667776
	Glossolepis dorityi	ZSM (IS50)*	Lake Nenggwambu, West Papua, Indonesia (AS)	KJ667878	KJ667783
	Glossolepis maculosus	ZSM 34147	Aquarium stock	KJ667879	KJ667784
	Glossolepis pseudoincisus	ZSM (IS52)*	Tami River System, Lake Yaniruk, West Papua, Indonesia (AS)	KJ667880	KJ667785
	Glossolepis romuensis	ZSM (IS51)*	Ramu River System, Papua New Guinea (AS)	KJ667881	KJ667786
	Melanotaenia affinis "Lae"	ZSM 34138	Aquarium stock	KJ667888	KJ667789
	Melanotaenia affinis "Pagwi"	ZSM 34146	Aquarium stock	KJ667889	KJ667790
	Melanotaenia angfa	ZSM 34148	Aquarium stock	KJ667890	KJ667791
	Melanotaenia angfa	ZSM 34149	Aquarium stock	KJ667891	KJ667792
	Melanotaenia australis	ZSM (IS57)*	Northern Australia (AS)	KJ667892	KJ667793
	Melanotaenia batanta	ZSM 34535	Batanta Island, West Papua, Indonesia (AS)	KJ667893	KJ667794
	Melanotaenia boesemani	ZSM 34153	Lake Aitinyo, West Papua, Indonesia (AS)	–	KJ667795
	Melanotaenia cf. eachamensis	ZSM 34158	upper Barron River, Queensland, Australia (AS)	KJ667894	KJ667796
	Melanotaenia douboulayi	ZSM (IS54)*	Eastern Australia (AS)	KJ667895	KJ667797
	Melanotaenia goldiei "Tapini"	ZSM 34174	Aquarium stock	KJ667896	KJ667798
	Melanotaenia herbertaxelrodi	ZSM (IS53)*	Lake Tebera, Papua New Guinea (AS)	KJ667897	KJ667799
	Melanotaenia lacustris	AP004419	Lake Kutubu, Papua New Guinea (AS)	–	(Miya et al., 2003)
	Melanotaenia lacustris	ZSM (IS58)*	Lake Kutubu, Papua New Guinea (AS)	KJ667898	KJ667800
	Melanotaenia macchullochi	ZSM (IS61)*	Eastern Australia (AS)	KJ667899	KJ667801
	Melanotaenia papuae	ZSM 34151	Aquarium stock	KJ667900	KJ667802
	Melanotaenia praecox	ZSM (IS55)*	Mamberamo River System, West Papua, Indonesia (AS)	KJ667901	KJ667803

Table continued

	Melanotaenia sexlineata "Kiunga"	ZSM 34155	Aquarium stock	KJ667903	KJ667805
	Melanotaenia sexlineata "Tabubil"	ZSM 34154	Aquarium stock	KJ667902	KJ667804
	Melanotaenia splendida inornata	ZSM 34167	Adelaide River, Northern Territory, Australia (AS)	KJ667908	KJ667810
	Melanotaenia splendida inornata	ZSM 34161	Blyth River, Northern Territory, Australia (AS)	KJ667904	KJ667806
	Melanotaenia splendida inornata	ZSM 34165	Flat Rock Creek, New South Wales, Australia (AS)	KJ667906	KJ667808
	Melanotaenia splendida inornata	ZSM 34163	Goyder Creek, Northern Territory, Australia (AS)	KJ667905	KJ667807
	Melanotaenia splendida inornata	ZSM 34166	Mann River, New South Wales, Australia (AS)	KJ667907	KJ667809
	Melanotaenia splendida splendida	ZSM 34160	Deepwater Creek, Queenslands, Australia (AS)	KJ667909	KJ667811
	Melanotaenia trifasciata "Hapgood River"	ZSM 34180	Aquarium stock	KJ667911	KJ667813
	Melanotaenia trifasciata	ZSM 34182	Pappan Creek, Queensland, Australia (AS)	KJ667912	KJ667814
	Melanotaenia trifasciata	ZSM 34157	Wonga Creek, Norther Territory, Australia (AS)	KJ667910	KJ667812
	Rhadinocentrus ornatus	ZSM 34156	Coolum Creek, Queensland, Australia (AS)	KJ667924	KJ667826
Pseudomugilidae	*Pseudomugil furcatus*	ZSM (IS68)*	Aquarium stock	KJ667918	KJ667820
	Pseudomugil pellucidus	ZSM (IS63)*	Aquarium stock	KJ667919	KJ667821
	Pseudomugil reticulatus	ZSM 34170	Aquarium stock	KJ667920	KJ667822
	Pseudomugil signifer	ZSM (IS66)*	Aquarium stock	KJ667921	KJ667823
Telmatherinidae	*Kalyptatherina helodes*	MZB 17152-1	Batanta, west of New Guinea	KJ667883	–
	Kalyptatherina helodes	MZB 17152-2	Batanta, west of New Guinea	KJ667884	–
	Kalyptatherina helodes	MZB 22095-1	Misool, west of New Guinea	KJ667885	–
	Kalyptatherina helodes	MZB 22095-2	Misool, west of New Guinea	KJ667886	–
	Marosatherina ladigesi	ZSM 34735	Maros karst, SW Sulawesi, Indonesia	KJ667887	KJ667788
	Paratherina cf. *labiosa*	ZSM 33054	Lake Mahalona, Sulawesi, Indonesia	KJ667913	KJ667815
	Paratherina cf. *labiosa*	ZSM 33055	Lake Mahalona, Sulawesi, Indonesia	KJ667914	KJ667816
	Paratherina cf. *striata*	ZSM 33057	Lake Mahalona, Sulawesi, Indonesia	KJ667915	KJ667817
	Paratherina cf. *striata*	ZSM 33061	Lake Towuti, Sulawesi, Indonesia	KJ667916	KJ667818
	Paratherina striata	ZSM 33062	Lake Mahalona, Sulawesi, Indonesia	KJ667917	KJ667819
	Telmatherina sarasinorum	ZSM 33115	Lake Matano, Sulawesi, Indonesia	KJ667957	KJ667859
	Telmatherina sarasinorum	ZSM 33116	Lake Matano, Sulawesi, Indonesia	KJ667954	KJ667856
	Telmatherina sarasinorum	ZSM 33117	Lake Matano, Sulawesi, Indonesia	KJ667955	KJ667857
	Telmatherina sarasinorum	ZSM 33127	Lake Matano, Sulawesi, Indonesia	KJ667956	KJ667858
	Telmatherina abendanoni	ZSM 33118	Lake Matano, Sulawesi, Indonesia	KJ667926	KJ667828
	Telmatherina abendanoni	ZSM (T26)*	Lake Matano, Sulawesi, Indonesia	KJ667925	KJ667827
	Telmatherina antoniae	ZSM 33089	Lake Matano, Sulawesi, Indonesia	KJ667927	KJ667829
	Telmatherina antoniae	ZSM 33090	Lake Matano, Sulawesi, Indonesia	KJ667928	KJ667830
	Telmatherina antoniae	ZSM 33092	Lake Matano, Sulawesi, Indonesia	KJ667929	KJ667831
	Telmatherina antoniae	ZSM 33093	Lake Matano, Sulawesi, Indonesia	KJ667930	KJ667832
	Telmatherina antoniae	ZSM 33100	Lake Matano, Sulawesi, Indonesia	KJ667931	KJ667833
	Telmatherina bonti	ZSM 33067	Stream close by Malili, Central Sulawesi, Indonesia	KJ667934	KJ667836

Table continued

	Telmatherina bonti	ZSM 33068	Stream close by Malili, Central Sulawesi, Indonesia	KJ667935	KJ667837
	Telmatherina bonti	ZSM 33070	Tominanga River (connecting Mahalona-Towuti), Sulawesi, Indonesia	KJ667936	KJ667838
	Telmatherina bonti	ZSM 33128	Nuha, Lake Matano, Sulawesi, Indonesia	KJ667937	KJ667839
	Telmatherina bonti	ZSM 33129	Nuha, Lake Matano, Sulawesi, Indonesia	KJ667938	KJ667840
	Telmatherina bonti	ZSM 33160	Tominanga River (connecting Mahalona-Towuti), Sulawesi, Indonesia	KJ667939	KJ667841
	Telmatherina bonti	ZSM (T85)*	Lake Towuti, Sulawesi, Indonesia	KJ667932	KJ667834
	Telmatherina bonti	ZSM (T86)*	Lake Towuti, Sulawesi, Indonesia	KJ667933	KJ667835
	Telmatherina celebensis	ZSM 33071	Lake Mahalona, Sulawesi, Indonesia	KJ667940	KJ667842
	Telmatherina celebensis	ZSM 33072	Lake Towuti, Sulawesi, Indonesia	KJ667941	KJ667843
	Telmatherina celebensis	ZSM 33078	Lake Mahalona, Sulawesi, Indonesia	KJ667942	KJ667844
	Telmatherina celebensis	ZSM 33079	Lake Towuti, Sulawesi, Indonesia	KJ667943	KJ667845
	Telmatherina cf. *obscura*	ZSM 33126	Lake Matano, Sulawesi, Indonesia	KJ667944	KJ667846
	Telmatherina cf. *wahjui*	ZSM (T53)*	Petea River (connecting Matano-Mahalona), Sulawesi, Indonesia	KJ667945	KJ667847
	Telmatherina cf. *wahjui*	ZSM (T54)*	Petea River (connecting Matano-Mahalona), Sulawesi, Indonesia	KJ667946	KJ667848
	Telmatherina opudi	ZSM 33120	Lake Matano, Sulawesi, Indonesia	KJ667947	KJ667849
	Telmatherina opudi	ZSM 33153	Lake Matano, Sulawesi, Indonesia	KJ667948	KJ667850
	Telmatherina opudi	ZSM 33154	Lake Matano, Sulawesi, Indonesia	KJ667949	KJ667851
	Telmatherina opudi	ZSM 33157	Lake Matano, Sulawesi, Indonesia	KJ667950	KJ667852
	Telmatherina prognatha	ZSM 33097	Lake Matano, Sulawesi, Indonesia	KJ667951	KJ667853
	Telmatherina prognatha	ZSM 33106	Lake Matano, Sulawesi, Indonesia	KJ667952	KJ667854
	Telmatherina prognatha	ZSM 33110	Lake Matano, Sulawesi, Indonesia	KJ667953	KJ667855
	Telmatherina sp.	ZSM 33121	Lake Matano, Sulawesi, Indonesia	KJ667958	KJ667860
	Telmatherina sp.	ZSM 33122	Lake Matano, Sulawesi, Indonesia	KJ667959	KJ667861
	Telmatherina wahjui	ZSM 33151	Lake Matano, Sulawesi, Indonesia	KJ667961	KJ667863
	Telmatherina wahjui	ZSM (T33)*	Lake Matano, Sulawesi, Indonesia	KJ667960	KJ667862
	Tominanga sp.	ZSM 33133	Lake Mahalona, Sulawesi, Indonesia	KJ667962	KJ667864
	Tominanga sp.	ZSM 33134	Lake Mahalona, Sulawesi, Indonesia	KJ667963	KJ667865

[1] MZB = Museum Zoologicum Bogoriense, Bogor; ZSM = Zoologische Staatssammlung München (Bavarian State Collection of Zoology, Munich); * = no voucher material available.

Supplementary Information – Chapter 4

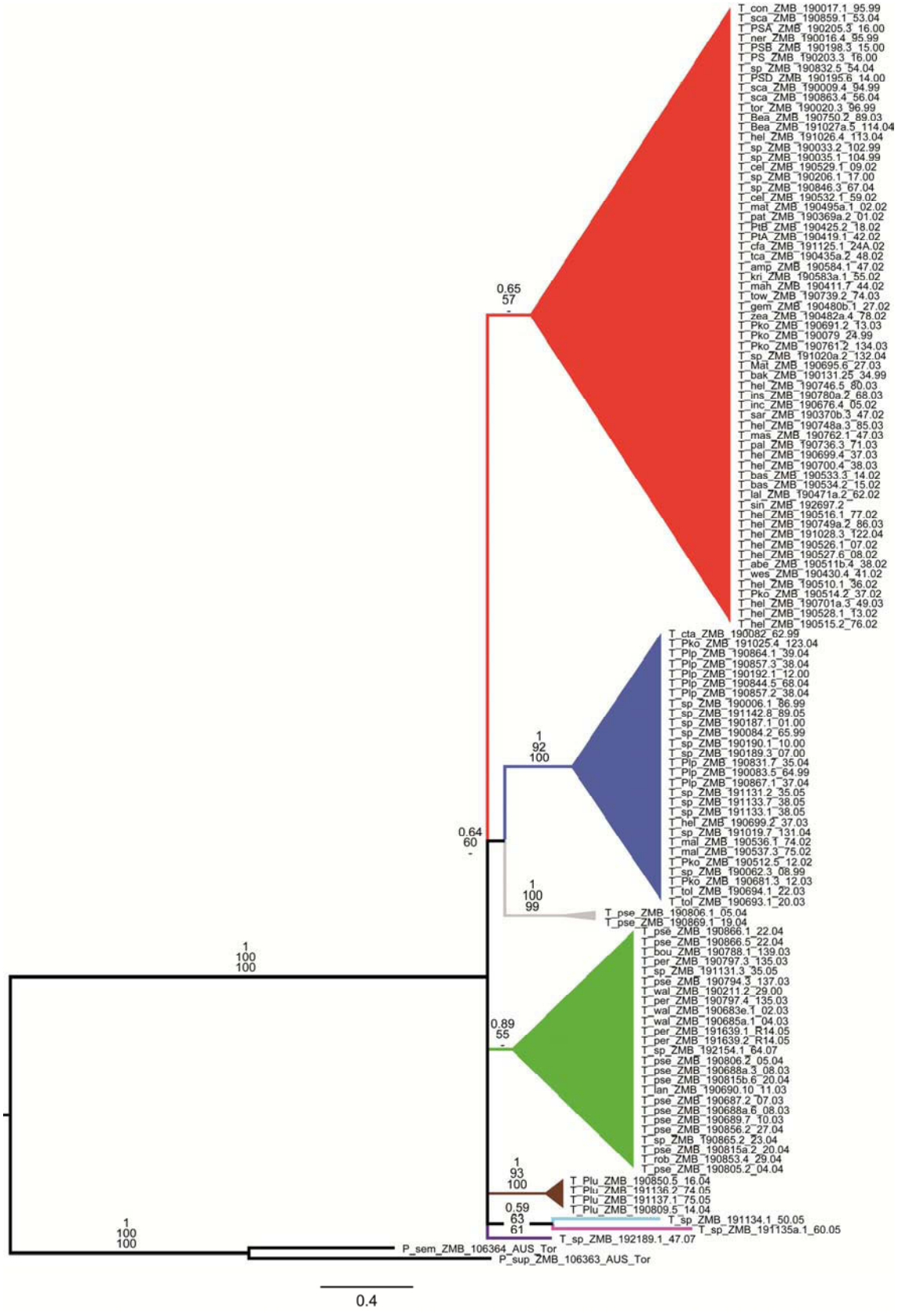


Figure S1. Molecular phylogeny of Tylomelania based on COI only (BI phylogram); see text for details. Numbers on branches show node support; BI posterior probability (top), ML (centre), and MP (bottom) bootstrap values. Colours for clades correspond to those used in Figure 3.

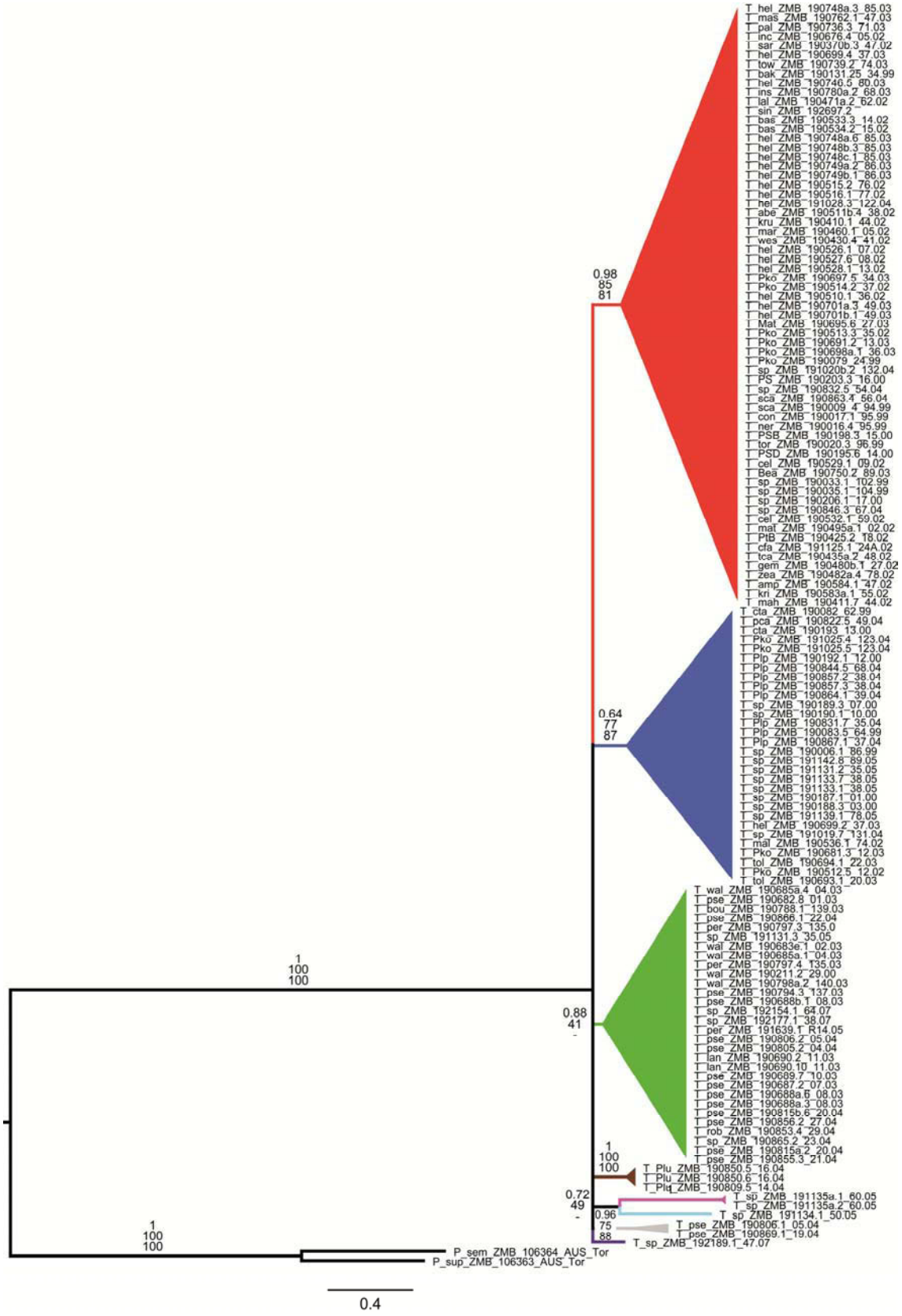


Figure S2. Molecular phylogeny of Tylomelania based on 16S only (BI phylogram); see text for details. Numbers on branches show node support; BI posterior probability (top), ML (centre), and MP (bottom) bootstrap values. Colours for clades correspond to those used in Figure 3.

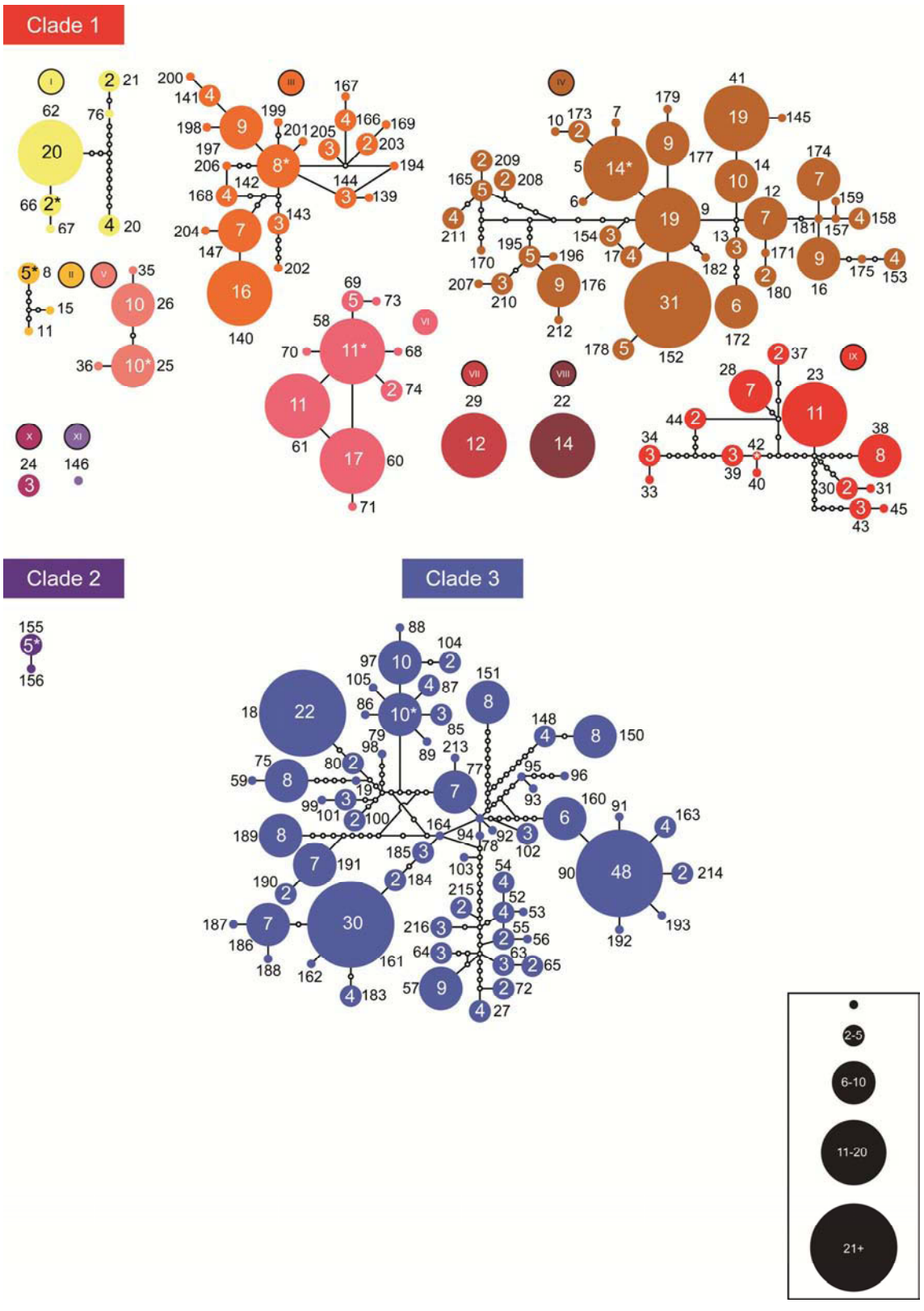


Figure S3. Haplotype networks for COI (part 1); see text for details. Colours for clades correspond to those used in Figure 3.

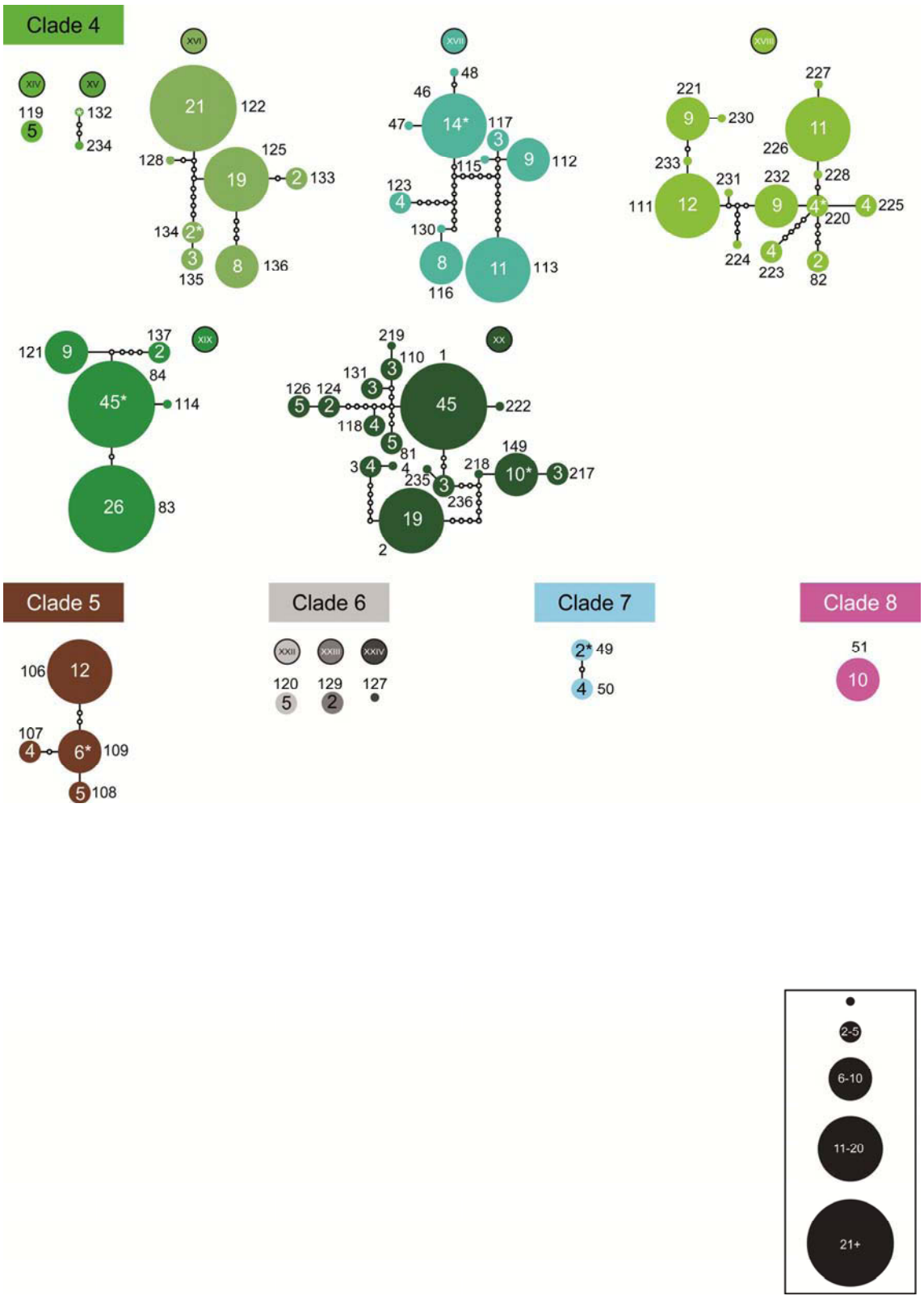


Figure S4. Haplotype networks for COI (part 2); see text for details. Colours for clades correspond to those used in Figure 3.

Table S1. List of sequenced specimens and sample provenience. The numbers listed under ZMB. Moll are the accession numbers of the malacological collection in the Museum für Naturkunde Berlin. This table lists all localities but only shows the accession numbers for sequences used for the tree reconstructions (Figures 1-3, S1, S2). Accession numbers for sequences that were only used for the computation of the haplotype networks (Fig. 3, S3, S4) are provided in Table S2, the respective samples are indicated by 'HN' in the COI sequence accession no. column. The letters in the haplotypes columns (A…Z, a…k) indicate specimens sharing the same haplotype, and bold type indicates the specimen shown in the tree(s). 1 = COI, 2 = 16S, 3 = concatenated dataset. All samples/sequences without an entry in the source column have been sequenced for this study.

Taxon	ZMB Moll.	Locality data	GenBank accession no.		Haplotypes			Source
			COI	16S	1	2	3	
Pseudopotamis		Australia: Queensland: Torres Strait Islands						
semoni	106363	Hammond Island, creek on footpath to old gold mine; 10 33.04'S, 142 12.76'E	AY242967	AY242968				Glaubrecht & Rintelen 2003
supralirata	106364	Prince of Wales Island, creek and pond at waterfall; 10 37.63'S, 142 11.53'E	AY242969	AY242970				Glaubrecht & Rintelen 2003
Tylomelania		Indonesia: Sulawesi						
abendanoni	190511b	South Sulawesi: Lake Lontoa: W shore	KJ850747	KJ850589	**A**			
amphiderita	190584	South Sulawesi: Lake Towuti: Loeha Island, W shore; 2 45.48'S, 121 30.9'E	KJ850748	KJ850590				
bakara	190131	South Sulawesi: Lake Towuti: NW shore, Cape Bakara; 2 40.76'S, 121 26.09'E	AY311974	AY311821				Rintelen et al. 2004
baskasti	190533	South Sulawesi: Larona catchment: Larona River; 2 40.36'S, 121 9.79'E	EF140859	EF140843				Rintelen et al. 2004
	190534	South Sulawesi: Larona catchment: Larona River; 2 40.55'S, 121 9.54'E	KJ850749	EU881926				this study / Rintelen & Glaubrecht 2008
carota	190082	South Sulawesi: Kalaena catchment: Kalaena River, upper reaches; 2 26.2'S, 120 48.53'E	AY311958	AY311823	**B**			Rintelen et al. 2004
	190193	South Sulawesi: Kalaena catchment: Stream at road Wotu – Pendolo; 2 22.93'S, 120 47.9'E	AY311957	AY311824	B			
celebicola	190529	Central Sulawesi: Tomori area: Lintumewure River, N of Nuha; 2 25.36'S, 121 21.43'E	KJ850750	KJ850591	**C**	**A**	**A**	
	190530	Central Sulawesi: Tomori area: Poaha River, N of Nuha; 2 22,055'S, 121 18,119'E	EF140864	EF140848	D	B	B	Rintelen et al. 2004
	190531	Sulawesi: Central Sulawesi: Tomori area: Lintumewure River, N of Nuha; 2 23.52'S, 121 19.66'E	KJ850751	KJ850592	C	A	A	
	190532	Central Sulawesi: Tomori area: Stream N of Nuha; 2 21.82'S, 121 18.05'E	EF140850	EF140833	**D**	**B**	**B**	Rintelen et al. 2004
			KJ850752	KJ850593	D	B	B	
	191094a	Central Sulawesi: Tomori area: Stream at road Beteleme – Nuha;	HN	-				
	191094b		HN	-				
	191109	Central Sulawesi: Tomori area: Stream at road Beteleme – Nuha; 2 13.52'S, 121 16.03'E	HN	-				
	191110	Central Sulawesi: Tomori area: Stream N of Wawopada, W of Beteleme; 2 09.05'S, 121 12.60'E	HN	-				
	191607	Central Sulawesi: Tomori area: Rumpi River, at road Nuha – Beteleme; 2 20.52'S, 121 18.06'E	HN	-				
	191608a-c	Central Sulawesi: Tomori area: Stream at road Beteleme – Nuha; 2 20.20'S, 121 18.16'E	HN	-				
	191609	Central Sulawesi: Tomori area: Stream at road Beteleme – Nuha; 2 15.37'S, 121 16.86'E	HN	-				
	191610	Central Sulawesi: Tomori area: River in Tingkeao, at road Beteleme – Tomata; 2 08.52'S, 121 14.73'E	HN	-				
confusa	191125	South Sulawesi: Lake Mahalona: SE shore, c. 30m W of Tominanga River outflow; 0 36,86'S, 121 30,9'E	KJ850753	KJ850594	**E**	**C**	**C**	
connectens	190017	Central Sulawesi: Poso River: Sulewana rapids; 1 38,9'S, 120 39,26'E	AY311961	AY311826				Rintelen et al. 2004
gemmifera	190480b	South Sulawesi: Lake Matano: N shore; 2 29,049'S, 121 25,034'E	KJ850754	KJ850595				
helmuti	190209	South Sulawesi: Towuti catchment: Stream at road Wawondula - Timampu; 2 38,46'S, 121 22,74'E	HN	-				
	190510		EF140851	EF140834				Rintelen et al. 2004
	190515	Southeast Sulawesi: Stream, just S of track Lengkobale - Parunela; 2 56,66'S, 121 30,81'E	EF140852	EF140835				Rintelen et al. 2004
	190516	Southeast Sulawesi: Tributary to Parudongka River, SE of Parunela; 2 56,74'S, 121 34,11'E	EF140853	EF140836	**F**			
	190525	South Sulawesi: Mahalona catchment: Tributary of Ponsoa River; 2 32,444'S, 121 33,345'E	HN	-				

Table continued

	190526	South Sulawesi: Mahalona catchment: Tributary of Ponsoa River; 2 32,33'S, 121 32,23'E	EF140854	EF140838				Rintelen et al. 2004
	190527	South Sulawesi: Petea catchment: Tributary of Petea River; 2 31,8'S, 121 29,99'E	EF140855	EF140839				Rintelen et al. 2004
	190528	South Sulawesi: Petea catchment: Tributary of Petea River; 2 31,79'S, 121 28,81'E	KJ850755	KJ850596	**G**	**D**		
	190699	South Sulawesi: Larona catchment: Kondube River, at road N of Matompi, nr Timampu; 2 38,36'S, 121 24,98'E	KJ850756	KJ850597				
			KJ850757	KJ850598		**E**		
	190700	South Sulawesi: Larona catchment: Stream at road Wawondula - Timampu; 2 38,72'S, 121 23,48'E	KJ850758	KJ85059		E		
	190701a	South Sulawesi: Towuti catchment: Bombongan River; 2 52,38'S, 121 22,54'E	KJ850759	KJ850600				
	190701b		EF140856	EF140840	G			Rintelen et al. 2004
	190746	South Sulawesi: Tominanga catchment: N Tributary of Lampesue River; 2 35,61'S, 121 36,98'E	KJ850760	KJ850601				
	190747	South Sulawesi: Tominanga catchment: N tributary of Lampesue River; 2 35,47'S, 121 35,66'E	KJ850761	KJ850602	G	D		
	190748a	South Sulawesi: Towuti catchment: Lemolemo River; 2 42,62'S, 121 40,99'E	KJ850762	KJ850603	**F**			
			KJ850763	KJ850604				
	190748b		EF140857	EF140841	F			Rintelen et al. 2004
	190748c		KJ850764	KJ850605	F			
	190749a	South Sulawesi: Towuti catchment: Lengkona River; 2 40,82'S, 121 41,77'E	KJ850765	KJ850606	**H**			
	190749b		EF140858	EF140842	H			Rintelen et al. 2004
	191026	South Sulawesi: Towuti catchment: Stream at road Beau - Masiku, tributary of Tokalimbu River; 2 50,732'S, 121 35,445'E	KJ850766	KJ850607		F		
	191028	South Sulawesi: Towuti catchment: Kakau River; 2 53,594'S, 121 31,458'E	KJ850767	KJ850608				
inconspicua	190676	South Sulawesi: Lake Mahalona: SW shore; 2 36,64'S, 121 28,54'E	KJ850768	KJ850609				
insulaesacrae	190780a	South Sulawesi: Lake Towuti: N shore, bay E of Cape Bintu; 2 39,48'S, 121 33,25'E	KJ850769	KJ850610				
kristinae	190583a	South Sulawesi: Lake Towuti: Loeha Island, N shore; 2 45,67'S, 121 33,49'E	KJ850770	KJ850611				
kruimeli	190410	South Sulawesi: Lake Mahalona: W shore; 2 35,88'S, 121 28,36'E	KJ850771	KJ850612	A			
lalemae	190471a	South Sulawesi: Lake Towuti: W shore, entrance of outlet bay, Cape Larona; 2 48,5'S, 121 25,04'E	KJ850772	KJ850613				
mahalonensis	190411	South Sulawesi: Lake Mahalona: W shore; 2 35,88'S, 121 28,36'E	KJ850773	KJ850614				
masapensis	190762	South Sulawesi: Lake Masapi: S shore; 2 50,84'S, 121 21,09'E	KJ850774	KJ850615				
marwotoae	190460	South Sulawesi: Lake Mahalona: SW shore; 2 36,64'S, 121 28,54'E	KJ850775	KJ850616	A			
matannensis	190495a	South Sulawesi: Lake Matano: S shore; 2 30,2'S, 121 19,37'E	KJ850776	KJ850617		**G**		
neritiformis	190016	Central Sulawesi: Poso River: Sulewana rapids; 1 38,9'S, 120 39,26'E	AY242959	AY242960		**H**		Rintelen et al. 2004
palicolarum	190736	South Sulawesi: Lake Towuti: E shore, off Beau village; 2 48,99'S, 121 33,64'E	KJ850777	KJ850618				
patriarchalis	190369a	South Sulawesi: Lake Matano: S shore, W of islets; 2 28,462's, 121 15,584'E	KJ850778	KJ850619		G		
perconica	190822	South Sulawesi: Kalaena catchment: Tributary of Tomoni River, W of Mangkutana; 2 29,207'S, 120 45,726'E	KJ850779	KJ850620	B			
perfecta	190797	South Sulawesi: Maros carst: Bantimurung River, above waterfall; 5 0,96'S, 119 40,92'E	KJ850780	KJ850621				
			KJ850781	KJ850622				
	191639	South Sulawesi: 166m; 4 58,432'S, 120 1,8'E	KJ850782	KJ850623	**J**	**J**	**D**	
			KJ850783	KJ850624	**K**	J	**E**	
robusta	190811	South Sulawesi: Langgara River (tributary of Dua River), Maroangin, off road Rappang - Enrekang; 3 44,414'S, 119 52,398'E	KJ850784	KJ850625	L	K	F	
	190845	South Sulawesi: Bilokka River, W of Bilokka, S of Amparita; 4 3,215'S, 119 49,1'E	KJ850785	KJ850626	L	K	F	
	190853	South Sulawesi: Tarrak River, S of Enrekang (=66-99); 3 41,762'S, 119 47,388'E	KJ850786	KJ850627				
	190865	South Sulawesi: Kupa, stream E of road Makassar - Parepare; 4 7,567'S, 119 37,426'E	KJ850787	KJ850628	**M**	**L**	**G**	
			KJ850788	KJ850629	M	L	G	

Table continued

	192152	South Sulawesi: River at road Camba - Watampone; 4 53.139'S, 119 53.527'E	HN	-				
	192154	South Sulawesi: River at Kampungbaru, W-bound dead-end road off road Maros - Soppeng; 4 39,334'S, 119 53,503'E	KJ850789	KJ850630				
	192155	South Sulawesi: Rumpia River, S of road Pekkae - Soppeng; 4 35,891'S, 119 41,097'E	KJ850790	KJ850631	N	M	H	
	193903	South Sulawesi: Pucue River, Datae, road Parapare - Sidenreng; 3 54,893'S, 119 42,571'E	HN	-				
sarasinorum	190370b	South Sulawesi: Lake Towuti: Loeha Island, W shore; 2 45,48'S, 121 30,9'E	KJ850791	KJ850632				
scalariopsis	190008	Central Sulawesi: Poso catchment: Saluopa waterfall; 1 44,95'S, 120 32,21'E	AY312007	AY311892	O	N	J	Rintelen et al. 2004
	190009	Central Sulawesi: Poso catchment: Stream c. 100m below Salopa waterfall; 1 44,9'S, 120 32,2'E	AY311962	AY311831				Rintelen et al. 2004
	190859	Central Sulawesi: Poso catchment: Creek in village, tributary of Poso River; 1 39,34'S, 120 39,957'E	KJ850792	KJ850633		H		
	190863	Central Sulawesi: Poso catchment: Saluopa waterfall; 1 44,95'S, 120 32,21'E	KJ850793	KJ850634	**O**	N	**J**	
	191603	Central Sulawesi: Poso catchment: Stream, N tributary of Lake Poso; 1 47.439'S, 120 30.799'E	HN	-				
sinabartfeldi	192697	South Sulawesi: Larona catchment: Larona River, at road Malili - Soroako; 2 40,29'S, 121 9,81'E	KJ850794	EU881924				this study / Rintelen & Glaubrecht 2008
spec. 1	190062	South Sulawesi: Matano catchment: Stream at SW shore of Lake Matano, c. 200-500m above mouth; 2 27,77'S, 121 13,35'E	AY312005	AY311888		O		Rintelen et al. 2004
	190750	South Sulawesi: Towuti catchment: Tokalimbo River; 2 48,89'S, 121 34,89'E	KJ850795	KJ850635	**P**	**F**	**K**	
			KJ850796	KJ850636	P	F	K	
	191027a	South Sulawesi: Towuti catchment: Masiku, Masiku River, tributary of Tokalimbu River;	KJ850797	KJ850637		F		
	191027b	2 51,988'S, 121 36,634'E	HN	-				
spec. 2	190690	South Sulawesi: Bone carst: Large carstic resurgence, Lanca, at road Baleleng - Tokaseng;	KJ850798	KJ850638	Q			Rintelen et al. 2004
		4 23,031'S, 120 13,965'E	KJ850799	KJ850639	**Q**	**P**	**L**	
			KJ850800	KJ850640	Q	P	L	
spec. 3	190079	South Sulawesi: Larona catchment: Kondara River, at road Malili - Soroako; 2 35,17'S, 121 17,34'E	KJ850801	KJ850641	**R**	**Q**		
	190512	South Sulawesi: Matano catchment: Stream at S shore of Lake Matano; 2 28,6'S, 121 15,586'E	EF140861	EF140845		O		Rintelen et al. 2004
	190513	South Sulawesi: Matano catchment: Tapulemo River, W of Pontada, Soroako; 2 31,39'S, 121 20,22'E	KJ850802	KJ850642	R			
	190514	South Sulawesi: Larona catchment: Kondara River, nr Leduledu; 2 35,17'S, 121 17,3'E	KJ850803	KJ850643	**S**	**R**	**M**	
	190681	South Sulawesi: Malili drainage: Pongkeru River, at road Malili - Tolala; 2 41,895'S, 121 8,359'E	KJ850804	KJ850644				
	190691	South Sulawesi: Larona catchment: River, Balambano village, at road Malili - Soroako; 2 38,142'S, 121 12,856'E	KJ850805	KJ850645		**S**		
	190692	South Sulawesi: Larona catchment: Patingko River, at road to Tabarano, off road Malili - Soroako N of Balambano; 2 38,222'S, 121 14,56'E	EF140862	EF140846	S	R	M	Rintelen et al. 2004
	190695	South Sulawesi: Matano catchment: Stream at S shore of Lake Matano; 2 29,68'S, 121 18,4'E	KJ850806	KJ850646	**T**			
	190696	South Sulawesi: Larona catchment: Patingko River, S of road Malili - Soroako; 2 37,17'S, 121 14,95'E	KJ850807	KJ850647	R	S		
	190697	South Sulawesi: Larona catchment: Patingko River, at road Malili - Soroako, in Togo village (=24-99); 2 36,53'S, 121 15,56'E	KJ850808	KJ850648	S			
	190698a	South Sulawesi: Larona catchment: Kondara River (tributary of Patingko River), at road Malili - Soroako, E of Leduledu; 2 35,11'S, 121 17,34'E	KJ850809	KJ850649	R			
	190698b	South Sulawesi: Larona catchment: Kondara River (tributary of Patingko River), at road Malili - Soroako, E of Leduledu; 2 35,11'S, 121 17,34'E	KJ850810	KJ850650	S	R	M	
	190761	South Sulawesi: Matano catchment: Tapulemo River, W of Pontada, Soroako; 2 31,39'S, 121 20,22'E	EF140863	EF140847	**U**	Q	**N**	Rintelen et al. 2004

Table continued

	191018	South Sulawesi: Kawata catchment: Kawata River, 50m W of road Ussu - Toletole, c. 2.5km S of Toletole; 2 33,008'S, 121 6,427'E	KJ850811	KJ850651	T	Q		
	191019	South Sulawesi: Cerekang catchment: Tributary of Wewu River, in Laroeha; 2 28,232'S, 121 4,122'E	KJ850812	KJ850652				
	191020a	South Sulawesi: Kawata catchment: Kawata River, at road Ussu - Toletole, c. 0.7km S of Toletole;	KJ850813	KJ850653		Q		
	191020b	2 31,95'S, 121 6,623'E	KJ850814	KJ850654	T			
	191024	South Sulawesi: Matano catchment: Soroako, stream W of Salonsa; 2 30,619'S, 121 19,785'E	KJ850815	KJ850655	U	Q	N	
	191025	South Sulawesi: Matano catchment: Soroako, stream W of Salonsa; 2 30,655'S, 121 19,671'E	KJ850816	KJ850656	**V**			
			KJ850817	KJ850657	V			
	191119	South Sulawesi: Larona catchment: Stream, tributary of Patingko River, nr Leduledu; 2 34.512'S, 121 16.524'E	HN	-				
	191606	South Sulawesi: Matano catchment: Soroako, Salonsa, tributary of Lake Matano, nr mouth (below 123-04); 2 30.375'S, 121 19.674'E	HN	-				
spec. 4	190083	South Sulawesi: Palopo plain: Stream along road Sabbang - Palopo, S of Sabbang; 2 36,68'S, 120 14,27'E	AY312001	AY311884				Rintelen et al. 2004
	190191	South Sulawesi: Palopo plain: Stream at road Palopo - Masamba; 2 39.98'S, 120 11.88'E	HN	-				
	190192	South Sulawesi: Palopo plain: Stream at road Palopo - Masamba; 2 34,36'S, 120 16,8'E	AY312000	AY311883				Rintelen et al. 2004
	190831	South Sulawesi: Palopo region: Malango River, Lebang, W of Palopo; 2 59,031'S, 120 9,879'E	KJ850818	KJ850658		**T**		
	190844	South Sulawesi: Palopo region: T. Nyarang River, btw Walenrang and Palopo; 2 57,756'S, 120 10,769'E	KJ850819	KJ850659				
	190857	South Sulawesi: Palopo plain: Kula River, Bone, N of Masamba; 2 32,635'S, 120 19,96'E	KJ850820	KJ850660				
			KJ850821	KJ850661				
	190864	South Sulawesi: Palopo plain: E of Patila, road Masamba - Wotu; 2 36,789'S, 120 34,383'E	KJ850822	KJ850662				
	190867	South Sulawesi: Palopo region: Latuppa, SW of Palopo; 3 1,515'S, 120 10,438'E	KJ850823	KJ850663				
	192134	South Sulawesi: Stream at road Masamba - Wotu; 2 34.356'S, 120 27.122'E	HN	-				
	193899	South Sulawesi: Palopo plain: Makawa River, tributary, road Palopo - Sabbang; 2 45.763'S, 120 8.503'E	HN	-				
	193902	South Sulawesi: Stream at road Masamba - Wotu; 2 34.356'S, 120 27.122'E	HN	-				
spec. 5	190809	Central Sulawesi: Palu valley: Sidaunta, at road Palu - Gimpu; 1 23,113'S, 119 58,563'E	KJ850824	KJ850664	**W**			
	190841	Central Sulawesi: Palu valley: Pakuli, at road Palu - Gimpu; 1 13,434'S, 119 56,861'E	KJ850825	KJ850665	W	U		
	190850	Central Sulawesi: Palu valley: Donggala Kecil, W of Palu; 0 53,943'S, 119 50,447'E	KJ850826	KJ850666	**X**			
			KJ850827	KJ850667	X	**U**		
	191136	Central Sulawesi: Palu valley: Ngangabomba River, S of Palu, Mantikole, Wera; 1 0,926'S, 119 51,72'E	KJ850828	KJ850668		U		
	191137	Central Sulawesi: Palu valley: Stream 15 km S of Palu, Mantikole; 1 1,174'S, 119 51,736'E	KJ850829	KJ850669		U		
spec. 6	190198	Central Sulawesi: Lake Poso: W shore, N of Bancea; 1 58,95'S, 121 35,11'E	AY311966	AY311834				Rintelen et al. 2004
spec. 7	190195	Central Sulawesi: Lake Poso: SW shore, E of Boe; 2 4,55'S, 120 39,45'E	AY311968	AY311838				Rintelen et al. 2004
spec. 8	115517	South Sulawesi: Maros carst: Assuloang Cave, main entrance; 4 55.746'S, 119 40.662'E	HN	-				
	190682	South Sulawesi: Maros carst: Tributary of Anda River, E of Celae, N of road; 4 55,216'S, 119 42,279'E	KJ850830	KJ850670	Y	**V**	**O**	Rintelen et al. 2004
	190683d		KJ850831	KJ850671	Y	V	O	
	190683g		KJ850832	KJ850672	Y	V	O	
	190683b		HN	-				
	190687	South Sulawesi: Bone carst: Stream, nr Walesso, at road Tacippi - Uloe; 4 27,65'S, 120 11,18'E	KJ850833	KJ850673				Rintelen et al. 2004

Table continued

190688a	South Sulawesi: Bone carst: Kacimpang River, in Watangulo, at road Mattirowalle - Baleleng; 4 24,385'S, 120 11,256'E	KJ850834	KJ850674	**Z**	**W**	**P**	
		KJ850835	KJ850675				
190688b		KJ850836	KJ850676	K			
		KJ850837	KJ850677	Z	J	E	
		KJ850838	KJ850678	K	W	P	
190689	South Sulawesi: Bone carst: Public swimming pool at stream, E of Citta, 100m S of road Baleleng - Tokaseng; 4 22,565'S, 120 11,572'E	KJ850839	KJ850679				
190790	South Sulawesi: Maros carst: Carstic resurgence, at cave entrance and river below; 4 54,17'S, 119 41,97'E	KJ850840	KJ850680	K	J	E	Rintelen et al. 2004
190793	South Sulawesi: Maros carst: River at road N of Balocci; 4 54,63'S, 119 41,6'E	KJ850841	KJ850681	K	J	E	Rintelen et al. 2004
		KJ850842	KJ850682	K	J	E	
190794	South Sulawesi: Maros carst: Carstic resurgence, in ricefields S of Balocci; 4 55,34'S, 119 40,95'E	KJ850843	KJ850683				
190796	South Sulawesi: Maros carst; Small carstic resurgence, N of road Tonasa - Balocci; 4 53,48'S, 119 38,36'E	HN	-				
190805	South Sulawesi: Carstic outflow, at road Pekkae - Ralla, W of Ralla; 4 32,354'S, 119 41,414'E	KJ850844	KJ850684				
190806	South Sulawesi: River at road Pekkae - Ralla, btw Ralla and Kessi; 4 29,914'S, 119 44,087'E	KJ850845	KJ850685				
		KJ850846	KJ850686	**N**	**M**	**H**	
190807	South Sulawesi: Stream at road Pekkae - Ralla, btw Ralla and Kessi; 4 29,926'S, 119 45,734'E	KJ850847	KJ850687	N	M	H	
190810	South Sulawesi: Langkemma River, NW of Langkemma; 4 28,702'S, 119 52,682'E	KJ850848	KJ850688	K	J	E	
190815a	South Sulawesi: E of Campalawampang, road Takkalasi - Watansoppeng; 4 20,369'S, 119 41,418'E	KJ850849	KJ850689	**a**	**X**	**Q**	
		KJ850850	KJ850690	a	X	Q	
190815b		KJ850851	KJ850691				
		KJ850852	KJ850692	a	X	Q	
190855	South Sulawesi: River at road Takkalasi - Watansoppeng, W of Watansoppeng; 4 20,199'S, 119 42,88'E	KJ850853	KJ850693	a			
190856	South Sulawesi: Karajae River, Salokarajae, road Rappang - Enrekang; 3 46,653'S, 119 51,114'E	KJ850854	KJ850694	L	**K**	**F**	
190862	South Sulawesi: Pucue River, Datae, road Parepare - Sidenreng; 3 54,893'S, 119 42,571'E	KJ850855	KJ850695	L	K	F	
		KJ850856	KJ850696	L	K	F	
190866	South Sulawesi: Pacekke, E of Kirukiru (road E of Mangkoso); 4 15,705'S, 119 42,062'E	KJ850857	KJ850697		**Y**		
		KJ850858	KJ850698		Y		
190869	South Sulawesi: E of Campalawampang, road Takkalasi - Watansoppeng; 4 19,919'S, 119 39,972'E	KJ850859	KJ850699	**b**	**Z**	**R**	
		KJ850860	KJ850700	b	Z	R	
191144a	South Sulawesi: River, tributary of Menraleng River, S of Takalala at road Takalala - Camba;	HN	-				
191144b	4 31.929'S, 119 57.784'E	HN	-				
191145	South Sulawesi: Stream nr Woddi, W of road Takalala - Pekkae; 4 28.195'S, 119 54.518'E	HN	-				
191146	South Sulawesi: Tributary of Labuajang River, at Langkemme, road Takalala - Pekkae; 4 29.587'S, 119 54.09'E	HN	-				
192146	South Sulawesi: River at dead-end road off road Sinjai - Matango; 4 55.97'S, 120 7.638'E	HN	-				
192164	South Sulawesi: Maros carst: River near karstic resurgence, N of Balocci; 4 54.152'S, 119 41.865'E	HN	-				
192171a	South Sulawesi: Maros carst: Karst river at dead-end road to Leang Londron; 4 51.929'S,	HN	-				
192171b	119 37.702'E	HN	-				
192183	South Sulawesi: Maros carst: River at road Maros - Camba, in Battunuang; 5 3.083'S, 119 43.109'E	HN	-				
192188	South Sulawesi: Maros carst: Stream N of road Maros - Camba; 5 2.569'S, 119 42.219'E	HN	-				

Table continued

spec. 9	190205	Central Sulawesi: Lake Poso: S shore, Pendolo, at Hotel Mulia; 2 3,91'S, 120 41,5'E	KJ850861	KJ850701	**c**	H	**S**	
	190861	Central Sulawesi: Lake Poso: E shore, Cape Watulunto; 2 0,825'S, 120 42,007'E	KJ850862	KJ850702	c	H	S	
spec. 10	190419	South Sulawesi: Petea River: c. 1km from mouth at Lake Mahalona; 2 34,27'S, 121 31,07'E	KJ850863	KJ850703		C		
spec. 11	190425	South Sulawesi: Petea River: c. 2 km from Lake Matano outflow; 2 32,67'S, 121 29,48'E	KJ850864	KJ850704				
spec. 12	190006	Southeast Sulawesi: Moramo waterfalls; 4 13,24'S, 122 44,66'E	AY312002	AY311891				Rintelen et al. 2004
	191142	Southeast Sulawesi: Stream at Panganjaya. Road SE of Punggaluku; 4 23,284'S, 122 35,553'E	KJ850865	KJ850705				
	191143	Southeast Sulawesi: Moramo waterfalls; 4 13.24'S, 122 44.66'E	HN	-				
spec. 13	190026	Central Sulawesi: Tomori area: Mawaro River at road Tomata - Beteleme; 2 0.25'S, 121 0.48'E	HN	-				
	190030	Central Sulawesi: Tomori area: River at road Tomata - Beteleme; 2 1.16'S, 121 1.94'E	HN	-				
	190031	Central Sulawesi: Tomori area: Matang River at Road Tomata - Beteleme; 2 5.93'S, 121 7.28'E	HN	-				
	190033	Central Sulawesi: Tomori area: Puawu River, waterfall at road Tomata - Beteleme; 2 5,93'S, 121 9,82'E	AY312012	AY311897				Rintelen et al. 2004
	190035	Central Sulawesi: Tomori area: Stream at road Tampira - Beteleme; 2 7,11'S, 121 20,16'E	AY312013	AY311898				
	191092a	Central Sulawesi: Tomori area: Ensa River, at road Tomata - Beteleme; 2 1.084'S, 121 4.321'E	HN	-				
	191092b		HN	-				
	191092c		HN	-				
	191093	Central Sulawesi: Tomori area: Ampere River, at road Tomata - Beteleme; 2 1.709'S, 121 5.38'E	HN	-				
	191112a	Central Sulawesi: Tomori area: River at road Tomata - Beteleme; 2 5.908'S, 121 7.273'E	HN	-				
	191112b		HN	-				
	191114	Central Sulawesi: Tomori area: Stream at road Tentena - Tomata (=98-99); 1 48.62'S, 120 46.43'E	HN	-				
	191115	Central Sulawesi: Tomori area: River at road Tomata - Beteleme (=99-99); 2 0.305'S, 121 0.456'E	HN	-				
	191116a	Central Sulawesi: Tomori area: River at road Tomata - Beteleme; 2 0.882'S, 121 2.019'E	HN	-				
	191116b		HN	-				
	191118a	Central Sulawesi: Tomori area: Stream at road Tomata - Beteleme; 2 1.059'S, 121 3.925'E	HN	-				
	191118b		HN	-				
	191120	Central Sulawesi: Tomori area: Stream, tributary of Puawu River, at road Tomata - Beteleme; 2 5.972'S, 121 9.672'E	HN	-				
	191121	Central Sulawesi: Tomori area: River at road Tomata - Beteleme; 2 4.911'S, 121 6.668'E	HN	-				
	191611a	Central Sulawesi: Tomori area: Puawu River, cascades, at road Beteleme - Tomata, W of Wawopada;	HN	-				
	191611b	2 9.314'S, 121 12.042'E	HN	-				
	191613a	Central Sulawesi: Tomori area: Puawu River, at road Beteleme - Tomata; 2 5.984'S, 121 7.325'E	HN	-				
	191613b		HN	-				
	192158	Central Sulawesi: Tomori area: Small river at road Taripa - Pendolo; 1 55.157'S, 120 47.697'E	HN	-				
	192160	Central Sulawesi: Tomori area: River at road Taripa - Pendolo; 1 59.163'S, 120 46.5'E	HN	-				
spec. 14	190084	South Sulawesi: Toraja land: Dolok River, at road Rantepao - Enrekang, N of Belajen; 3 18,14'S, 119 49,35'E	AY312017	AY311901		T		
	190189	South Sulawesi: Toraja land: Rantepao, Marante River; 2 57,42'S, 119 55,78'E	AY312015	AY311899	**d**			Rintelen et al. 2004
	190190	South Sulawesi: Toraja land: Tributary of Kambuna River, off road Rantepao - Palopo; 2 54,97'S, 119 59,26'E	AY312016	AY311900				Rintelen et al. 2004
	190840	South Sulawesi: Toraja land: Pakampan (tributary of Saa River, Noling drainage); 3 4,169'S, 119 58,638'E	KJ850866	KJ850706	d	T	**T**	
	190851	South Sulawesi: Toraja land: Marara River, W of Ledo (Sadang drainage); 3 1,993'S, 119 56,409'E	KJ850867	KJ850707	d	T	T	

Table continued

	190868	South Sulawesi: Toraja land: Lengko River (Noling drainage), N of Simparrun; 3 4,887'S, 119 58,738'E	KJ850868	KJ850708	d	T	T	
	191593a	South Sulawesi: Toraja land: Natural swimming pool "Tilangga"; 3 2.129'S, 119 53.233'E	HN	-				
	191593b		HN	-				
	191594	South Sulawesi: Toraja land: Sa'dan River, at road Rantepao - Makale; 3 3.595'S, 119 52.07'E	HN	-				
	191595	South Sulawesi: Toraja land: Tributary of Rembon River and small inflowing stream, at road Makale - Mamasa; 3 0.478'S, 119 42.648'E	HN	-				
	191596	South Sulawesi: Toraja land: Stream at road Makale - Mamasa; 3 1.748'S, 119 44.318'E	HN	-				
spec. 15	190187	Southeast Sulawesi: Amberewa River, at road Kolaka - Kendari; 3 58,03'S, 122 15,77'E	AY312003	AY311885	**e**			Rintelen et al. 2004
	190188	Southeast Sulawesi: Watudehoa River, at road Lambuya - Aopa; 4 2,96'S, 122 6,86'E	AY312004	AY311886	e			
	191138	Southeast Sulawesi: Stream in Labela, E of Pundidaha, road Kendari - Kolaka; 3 57.897'S, 122 20.739'E	HN	-				
	191139	Southeast Sulawesi: Humbuti River, NW of Abuki; 3 38,91'S, 121 52,38'E	KJ850869	KJ850709	e			
	191140	Southeast Sulawesi: Simbune River, c. 1km NE of Raterate, road Kendari - Kolaka; 4 2.326'S, 121 54.204'E	HN	-				
	191141a	Southeast Sulawesi: Poniponiki River, NW of Raterate, road Kendari - Kolaka; 4 2.115'S,	HN	-				
	191141b	121 52.816'E	HN	-				
spec. 16	190206	Central Sulawesi: Poso catchment: Stream at road Pendolo - Bancea; 2 2,37'S, 120 37,25'E	AY312009	AY311887				Rintelen et al. 2004
spec. 17	190812	Central Sulawesi: Poso River: Sulewana, rapids; 1 38,871'S, 120 39,279'E	KJ850870	KJ850710	c	H	S	
spec. 18	191131	South Sulawesi: Creek running from NE into Pamukulu River at road crossing; 5 24,422'S, 119 36,722'E	KJ850871 KJ850872	KJ850711 KJ850712				
	191132	South Sulawesi: Creek running into Pamukulu River, further W, tributary down to river; 5 24,116'S, 119 36,24'E	HN	-				
	191133	South Sulawesi: Eremerasa swimming pool, N of Bantaeng; 5 28,896'S, 120 0,313'E	KJ850873 KJ850874	KJ850713 KJ850714				
	192143	South Sulawesi: River at road Bulukumba - Sinjai; 5 19,659'S, 120 7,97'E	KJ850875	KJ850715	J	J	D	
	192177	South Sulawesi: Malawa spring, near Patuku; 5 3,512'S, 119 57,475'E	KJ850876	KJ850716	J			
	193900	South Sulawesi: River at road Sinjai - Manipi, c. 2 km before Manipi; 5 13.2'S, 120 0.49'E	HN	-				
	193901	South Sulawesi: Aparang River, at Road Bulukumba - Sinjai, N of Bikeru; 5 15.13'S, 119 10.14'E	HN	-				
spec. 19	190693	Southeast Sulawesi: Tolala Region: Tolala River, E of Tolala at road to Loka; 2 56,05'S, 121 6,96'E	KJ850877	KJ850717				
	190694	Southeast Sulawesi: Tolala Region: River at road Malili - Tolala; 2 55,21'S, 121 5,7'E	EF140865	EF140849				Rintelen et al. 2004
	190839	Southeast Sulawesi: Tolala Region: River at road Malili - Tolala; 2 55,21'S, 121 5,7'E	HN	-				
spec. 20	190207	Central Sulawesi: Poso catchment: Stream at road Bancea - Tentena; 1 49.01'S, 120 31.03'E	HN	-				
	190832	Central Sulawesi: Poso catchment: Creek, tributary of Poso River; 1 41,019'S, 120 39,496'E	KJ850878	KJ850718				
	190846	Central Sulawesi: Poso catchment: Salokuwa River, tributary of Kodina River, S of Pendolo; 2 9.399'S, 120 43.943'E	KJ850879	KJ850719				
	191597	Central Sulawesi: Poso catchment: Stream at road Tentena - Peura, E shore of Lake Poso; 1 49.983'S, 120 38.242'E	HN	-				
	191598	Central Sulawesi: Poso catchment: Small stream, road Tentena - Toaro; 1 44.079'S, 120 40.098'E	HN	-				
	191599	Central Sulawesi: Poso catchment: Small stream, road Tentena - Toaro; 1 44.234'S, 120 40.039'E	HN	-				
	191600	Central Sulawesi: Poso catchment: Kelei, c. 5km SE of Tentena, small stream, tributary of Wimbi River; 1 47.5'S, 120 41.413'	HN	-				

Table continued

	191601	Central Sulawesi: Poso catchment: Tributary of Kaiuku River, c. 30m from river, at road Tentena - Pendolo; 1 46.29'S, 120 42.98'E	HN	-				
	191602	Central Sulawesi: Poso catchment: Stream at road Poso - Tentena; 1 37.48'S, 120 42.131'E	HN	-				
	191604	Central Sulawesi: Poso catchment: Tonusu, c. 3km S of Tonusu towards Siuri; 1 48.95'S, 120 31.18'E	HN	-				
	191605	Central Sulawesi: Poso catchment: Uebangke River, N tributary of Lake Poso; 1 46.48'S, 120 35.61'E	HN	-				
spec. 21	191134	Central Sulawesi: Luwuk Peninsula: Biak River, above cascades, at road Luwuk - Pagimana, S of Salodik; 0 50,242'S, 122 52,394'E	KJ850880	KJ850720	**f**	**a**	**U**	
			KJ850881	KJ850721	f	a	U	
	191135a	Central Sulawesi: Luwuk Peninsula: Stream in mountains nr Sape, SW of Balantak, E coast of Balantak Peninsula; 0 56,682'S, 123 21,38'E	KJ850882	KJ850722	**g**			
			KJ850883	KJ850723	g			
spec. 22	192189	Central Sulawesi: River at road Kotaraya - Tolitoli; 0 39,807'N, 120 39,409'E	KJ850884	KJ850724	**h**	**b**	**V**	
			KJ850885	KJ850725	h	b	V	
spec. 23	190788	South Sulawesi: Maros carst: Carstic resurgence, at cave entrance and river below; 4 54,17'S, 119 41,97'E	KJ850886	KJ850726	**Y**	**c**	**W**	
			KJ850887	KJ850727	Y	c	W	
	190791	South Sulawesi: Maros carst: River at road N of Balocci; 4 54,63'S, 119 41,6'E	KJ850888	KJ850728	Y	c	W	
	192149	South Sulawesi: Maros carst: Karstic river at mouth of Leang Londron cave; 4 51.746'S, 119 38.047'E	HN	-				
	192162	South Sulawesi: Maros carst: River near karstic resurgence, N of Balocci; 4 54.152'S, 119 41.865'E	HN	-				
	192173	South Sulawesi: Maros carst: Karst river at dead-end road to Leang Londron; 4 51.929'S, 119 37.702'E	HN	-				
spec. 24	190536	Southeast Sulawesi: Malili catchment: Stream at road Malili - Tolala; 2 49,77'S, 121 3,67'E	EF140860	EF140844		**d**		Rintelen et al. 2004
	190537	Southeast Sulawesi: Malili catchment: Stream at road Malili - Tolala; 2 48,9'S, 121 3,62'E	KJ850889	KJ850729		d		
spec. 25	190032	Central Sulawesi: Tomori area: Matang River at Road Tomata - Beteleme; 2 5.93'S, 121 7.28'E	HN	-				
	191113	Central Sulawesi: Tomori area: River at road Tomata - Beteleme; 2 5.908'S, 121 7.273'E	HN	-				
	191612	Central Sulawesi: Tomori area: Puawu River, at road Beteleme - Tomata; 2 5.984'S, 121 7.325'E	HN	-				
spec. 26	195096	South Sulawesi: Maros carst: Sambueja Cave, river in front of cave; 5 3.75'S, 119 41.28'E	HN	-				
toradjarum	190020	Central Sulawesi: Lake Poso: Tentena, E shore, near outflow of Poso River; 1 45,92'S, 120 38,42'E	AY242961	AY242962				Rintelen et al. 2004
	190203	Central Sulawesi: Lake Poso: S shore, Pendolo, at Hotel Mulia; 2 3,91'S, 120 41,5'E	AY311967	AY311836				Rintelen et al. 2004
towutensis	190739	South Sulawesi: Lake Towuti: S shore, c. 2 km E of Cape Mea; 2 55,8'S, 121 26,92'E	KJ850890	KJ850730				
towutica	190435a	South Sulawesi: Lake Towuti: Loeha Island, W shore; 2 46,18'S, 121 31,3'E	KJ850891	KJ850731				
wallacei	115515	South Sulawesi: Maros carst: Rumbia Cave, pond in front of cave outflow; 5 2.185'S, 119 37.683'E	HN	-				
	115516	South Sulawesi: Maros carst: small stream below gallery entrance of Gua Assuloang; 4 55.579'S, 119 40.716'E	HN	-				
	190211	South Sulawesi: Lejang, public pool "Mattampa", at road Pare-Pare - Makassar, N of Pangkajene; 4 48,38'S, 119 32,66'E	AY312006	AY311903	**j**			
	190683e	South Sulawesi: Maros carst: Tributary of Anda River, E of Celae, N of road; 4 55,216'S,	KJ850892	KJ850732				
	190683a,c,f	119 42,279'E	HN	-				
	190685a	South Sulawesi: Maros carst: Anda River, W of Celae; 4 54,728'S, 119 40,632'E	KJ850893	KJ850733	**k**	**e**	**X**	
			KJ850894	KJ850734	Y	**f**	**Y**	
	190685b		KJ850895	KJ850735	k	e	X	Rintelen et al. 2004
	190685c		KJ850896	KJ850736	Y	f	W	
			KJ850897	KJ850737	Y	c	Y	

Table continued

	190686	South Sulawesi: Maros carst: Anda River, W of Celae; 4 54,728'S, 119 40,632'E	HN	-				
	190789	South Sulawesi: Maros carst: Carstic resurgence, at cave entrance and river below; 4 54,17'S, 119 41,97'E	KJ850898	KJ850738	Y	V	O	
	190792	South Sulawesi: Maros carst: River at road N of Balocci; 4 54,63'S, 119 41,6'E	KJ850899	KJ850739	Y	c	W	Rintelen et al. 2004
			KJ850900	KJ850740	K	J	E	
	190795	South Sulawesi: Maros carst: Small carstic resurgence, N of road Tonasa - Balocci; 4 53,48'S, 119 38,36'E	KJ850901	KJ850741	Y	c	W	
	190798a	South Sulawesi: Maros carst: Lejang, public pool "Mattampa", at road Pare-Pare - Makassar, N of	KJ850902	KJ850742	j	**g**	**Z**	
	190798b	Pangkajene (=29-00); 4 48,38'S, 119 32,66'E	KJ850903	KJ850743	j	g	Z	
	192148	Sulawesi: South Sulawesi: River at road Sinjai - Matango; 4 45 973'S, 120 4.357'E	HN	-				
	192150	South Sulawesi: Maros carst: Karstic river at mouth of Leang Londron cave; 4 51.746'S, 119 38.047'E	HN	-				
	192163	South Sulawesi: Maros carst: River near karstic resurgence, N of Balocci; 4 54.152'S, 119 41.865'E	HN	-				
	192167a,b	South Sulawesi: Maros carst: Concrete canal at dead-end road to Leang Londron; 4 52.259'S, 119 37.508'E	HN	-				
	192172a,b	South Sulawesi: Maros carst: Karst river at dead-end road to Leang Londron; 4 51.929'S, 119 37.702'E	HN	-				
	192182	South Sulawesi: Maros carst: River at road Maros - Camba, in Battunuang; 5 3.083'S, 119 43.109'E	HN	-				
	192185	South Sulawesi: Maros carst: Stream N of road Maros - Camba; 5 2.569'S, 119 42.219'E	HN	-				
wesseli	190430	South Sulawesi: Tominanga River: c. 4.5 km N of mouth at Lake Towuti; 2 37,79'S, 121 31,81'E	KJ850904	KJ850744				
wolterecki	190463	South Sulawesi: Lake Mahalona: S shore, c. 50m W of Tominanga River outflow; 2 36,86'S, 121 30,9'E	KJ850905	KJ850745	E	C	C	
zeamais	190482a	South Sulawesi: Lake Matano: N shore, E of Nuha; 2 27,342'S, 121 21,624'E	KJ850906	KJ850746				

Table S2. Haplotypes of *Tylomelania* (see Figures S3 & S4) and their assignment to species and museum vouchers. The numbers listed under ZMB Moll. are the museum accession numbers of the malacological collection of the Museum für Naturkunde Berlin and provide a link to the sample locality information in Table S1. GenBank accession numbers (for each haplotype) are for the specimen indicated in brackets.

Haplotype	N	Species	ZMB Moll.	Individuals	GenBank accession no.
1	45	*Tylomelania* spec. 23	190788	1-6,8	KJ850886 (190788.1)
		Tylomelania spec. 23	190791	1-3	
		Tylomelania spec. 23	192162	1-6	
		Tylomelania spec. 8	190682	1-8	
		Tylomelania spec. 8	190683b	1	
		Tylomelania spec. 8	190683d	3	
		Tylomelania spec. 8	190683g	1-3	
		Tylomelania spec. 8	190790	5	
		Tylomelania robusta	192154	3-5	
		Tylomelania wallacei	115516	4	
		Tylomelania wallacei	190685a	2,4	
		Tylomelania wallacei	190685b	2,3	
		Tylomelania wallacei	190685c	3,4	
		Tylomelania wallacei	190789	3	
		Tylomelania wallacei	190792	1	
		Tylomelania wallacei	190795	3	
		Tylomelania wallacei	192163	2,4	
2	19	*Tylomelania* spec. 23	192149	1,2,5-8	KJ933712 (192149.1)
		Tylomelania spec. 23	192173	1-3	
		Tylomelania spec. 8	192171a	1	
		Tylomelania spec. 8	192171b	1-4	
		Tylomelania wallacei	192150	1-3,8	
		Tylomelania wallacei	192172b	1	
3	4	*Tylomelania* spec. 23	192149	3,4	KJ933713 (192149.3)
		Tylomelania spec. 23	192173	5,6	
4	1	*Tylomelania* spec. 23	192173	4	KJ933714 (192173.4)
5	14	*Tylomelania celebicola*	190529	1,2,4-7	KJ850750 (190529.1)
		Tylomelania celebicola	190531	1,2	
		Tylomelania celebicola	191609	1	
		Tylomelania spec. 13	191093	1,3,4	
		Tylomelania celebicola	191094b	2	
		Tylomelania spec. 13	191120	8	
6	1	*Tylomelania celebicola*	190529	3	KJ933715 (190529.3)
7	1	*Tylomelania celebicola*	190529	8	KJ933716 (190529.8)
8	5	*Tylomelania celebicola*	190530	2	
		Tylomelania celebicola	190532	1-4	KJ850752 (190532.2)
9	19	*Tylomelania celebicola*	190531	3	KJ933717 (190531.3)
		Tylomelania celebicola	191607	1-5,7	
		Tylomelania celebicola	191608a	1-6	
		Tylomelania celebicola	191608b	1-5	
		Tylomelania celebicola	191608c	1	
10	1	*Tylomelania celebicola*	191109	1	KJ933718 (191109.1)
11	1	*Tylomelania celebicola*	191109	2	KJ933719 (191109.2)
12	7	*Tylomelania celebicola*	191110	1	KJ933720 (191110.1)
		Tylomelania spec. 13	191120	2,6	
		Tylomelania spec. 13	191611a	2	
		Tylomelania spec. 13	191611b	1	
		Tylomelania spec. 13	191613b	1,2	
13	3	*Tylomelania celebicola*	191110	2,4,5	KJ933721 (191110.2)
14	10	*Tylomelania celebicola*	191110	3	KJ933722 (191110.3)
		Tylomelania celebicola	191610	1-4,7-11	
15	1	*Tylomelania celebicola*	191607	6	KJ933723 (191607.6)
16	9	*Tylomelania celebicola*	191610	5	KJ933724 (191610.5)
		Tylomelania spec. 13	191092b	1	
		Tylomelania spec. 13	191116a	2,3	
		Tylomelania spec. 13	191116b	1	
		Tylomelania spec. 13	191118b	1,3,6	
		Tylomelania spec. 13	191611b	3	
17	4	*Tylomelania celebicola*	191610	6	KJ933725 (191610.6)
		Tylomelania spec. 13	191118a	1,4	
		Tylomelania spec. 13	191118b	2	
18	22	*Tylomelania carota*	190082	1-6	AY311958 (190082.1)
		Tylomelania carota	190193	1,3-5	
		Tylomelania spec. 3	191018	2	
		Tylomelania spec. 3	191019	2-6,8	
		Tylomelania spec. 3	191020b	1	
		Tylomelania perconica	190822	2,3,5,6	
19	1	*Tylomelania carota*	190193	2	AY311957 (190193.2)
20	4	*Tylomelania helmuti*	190209	21	
		Tylomelania helmuti	190510	1,2	EF140851 (190510.1)

Table continued

		Tylomelania helmuti	190699	3	
21	2	*Tylomelania helmuti*	190510	3,4	KJ933726 (190510.3)
22	14	*Tylomelania helmuti*	190515	1-12	EF140852 (190515.2)
		Tylomelania helmuti	190516	2,3	
23	11	*Tylomelania helmuti*	190516	1	EF140853 (190516.1)
		Tylomelania helmuti	190748a	1,2,4-7	
		Tylomelania helmuti	190748b	1,3	
		Tylomelania helmuti	190748c	1,2	
24	3	*Tylomelania helmuti*	190525	2	KJ933727 (190525.2)
		Tylomelania helmuti	190526	4,5	
25	10	*Tylomelania helmuti*	190525	6	
		Tylomelania helmuti	190526	1	EF140854 (190526.1)
		Tylomelania helmuti	190528	1,2	
		Tylomelania helmuti	190746	4	
		Tylomelania helmuti	190747	2,4-7	
26	10	*Tylomelania helmuti*	190527	1-6	EF140855 (190527.6)
		Tylomelania helmuti	190528	3-6	
27	4	*Tylomelania helmuti*	190699	1,2	KJ850756 (190699.2)
		Tylomelania spec. 3	190692	5	
		Tylomelania spec. 3	190698a	6	
28	7	*Tylomelania helmuti*	190700	1,3-8	KJ850758 (190700.4)
29	12	*Tylomelania helmuti*	190701a	1,3,5-10,12,13	
		Tylomelania helmuti	190701b	1,2	EF140856 (190701b.1)
30	2	*Tylomelania helmuti*	190701a	2,11	KJ933728 (190701a.2)
31	1	*Tylomelania helmuti*	190701a	4	KJ933729 (190701a.4)
32	1	*Tylomelania helmuti*	190746	1	KJ933730 (190746.1)
33	3	*Tylomelania helmuti*	190746	2,3,5	KJ850760 (190746.5)
34	1	*Tylomelania helmuti*	190747	1	KJ933731 (190747.1)
35	1	*Tylomelania helmuti*	190747	3	KJ933732 (190747.3)
36	2	*Tylomelania helmuti*	190748a	3	KJ850762 (190748a.3)
		Tylomelania helmuti	190748b	2	
37	8	*Tylomelania helmuti*	190749a	1-6	KJ850765 (190749a.2)
		Tylomelania helmuti	190749b	1,2	
38	3	*Tylomelania helmuti*	191026	1,3,5	KJ933733 (191026.1)
39	1	*Tylomelania helmuti*	191026	2	KJ933734 (191026.2)
40	19	*Tylomelania helmuti*	191026	4	KJ850766 (191026.4)
		Tylomelania spec. 1	190750	1-6	
		Tylomelania spec. 1	191027a	1-4,6,7	
		Tylomelania spec. 1	191027b	1-6	
41	1	*Tylomelania helmuti*	191026	6	KJ933735 (191026.6)
42	3	*Tylomelania helmuti*	191028	1,4,6	KJ933736 (191028.1)
43	2	*Tylomelania helmuti*	191028	2,5	KJ933737 (191028.2)
44	1	*Tylomelania helmuti*	191028	3	KJ850767 (191028.3)
45	14	*Tylomelania* spec. 2	190690	1-7,9, 10,12-16	KJ850798 (190690.2)
46	1	*Tylomelania* spec. 2	190690	11	KJ933738 (190690.11)
47	1	*Tylomelania* spec. 2	190690	8	KJ933739 (190690.8)
48	2	*Tylomelania* spec. 21	191134	1,2	KJ850880 (191134.1)
49	4	*Tylomelania* spec. 21	191134	3-6	KJ933740 (191134.3)
50	10	*Tylomelania* spec. 21	191135a	1-10	KJ850882 (191135a.1)
51	4	*Tylomelania* spec. 24	190536	1-3,5	EF140860 (190536.1)
52	1	*Tylomelania* spec. 24	190536	4	KJ933741 (190536.4)
53	4	*Tylomelania* spec. 24	190536	6	
		Tylomelania spec. 24	190537	1-3	KJ850889 (190537.3)
54	2	*Tylomelania* spec. 24	190537	4,5	KJ933742 (190537.4)
55	1	*Tylomelania* spec. 24	190537	6	KJ933743 (190537.6)
56	9	*Tylomelania* spec. 3	190512	1,3-5	EF140861 (190512.5)
		Tylomelania spec. 1	190062	3,5-8	
57	11	*Tylomelania* spec. 3	190512	2	
		Tylomelania spec. 3	190697	3	
		Tylomelania spec. 3	191018	4	KJ850811 (191018.4)
		Tylomelania spec. 3	191020a	1	
		Tylomelania spec. 3	191020b	2	
		Tylomelania spec. 3	190695	1-6	
58	1	*Tylomelania* spec. 3	190512	6	KJ933744 (190512.6)
59	17	*Tylomelania* spec. 3	190513	1,3,4	KJ850802 (190513.3)
		Tylomelania spec. 3	190514	1	
		Tylomelania spec. 3	190691	1-6	
		Tylomelania spec. 3	190696	2,9	
		Tylomelania spec. 3	190698a	1,3,7	
		Tylomelania spec. 3	190761	3,6	
60	11	*Tylomelania* spec. 3	190513	2	
		Tylomelania spec. 3	190692	4	
		Tylomelania spec. 3	190696	1,4,7	
		Tylomelania spec. 3	190761	1,2,4	EF140863 (190761.2)
		Tylomelania spec. 3	191024	1-3	

Table continued

61	20	*Tylomelania* spec. 3	190514	2,6	KJ850803 (190514.2)
		Tylomelania spec. 3	190692	3	
		Tylomelania spec. 3	190696	5	
		Tylomelania spec. 3	190697	1,4,5	
		Tylomelania spec. 3	190698a	2,4,5,8	
		Tylomelania spec. 3	190698b	1,2	
		Tylomelania spec. 3	191119	1,2,4,5	
62	3	*Tylomelania* spec. 3	190681	1,2,6	KJ933745 (190681.1)
63	3	*Tylomelania* spec. 3	190681	3-5	KJ850804 (190681.3)
64	2	*Tylomelania* spec. 3	190681	7,8	KJ933746 (190681.7)
65	2	*Tylomelania* spec. 3	190692	1	KJ933747 (190692.1)
		Tylomelania spec. 3	190697	2	
66	1	*Tylomelania* spec. 3	190692	2	KJ933748 (190692.2)
67	1	*Tylomelania* spec. 3	190696	3	KJ933749 (190696.3)
68	5	*Tylomelania* spec. 3	190696	6	KJ933750 (190696.6)
		Tylomelania spec. 3	191018	1,3,5,6	
69	1	*Tylomelania* spec. 3	190696	8	KJ933751 (190696.8)
70	1	*Tylomelania* spec. 3	190761	5	KJ933752 (190761.5)
71	2	*Tylomelania* spec. 3	191019	1,7	KJ850812 (191019.7)
72	1	*Tylomelania* spec. 3	191020a	2	KJ850813 (191020a.2)
73	2	*Tylomelania* spec. 3	191020a	3,4	KJ933753 (191020a.3)
74	8	*Tylomelania* spec. 3	191025	1-6	KJ850816 (191025.4)
		Tylomelania spec. 3	191606	1,2	
75	1	*Tylomelania* spec. 3	191119	3	KJ933754 (191119.3)
76	7	*Tylomelania perconica*	190083	6-8,10	KJ933755 (190083.6)
		Tylomelania spec. 4	190844	1	
		Tylomelania sp.	193899	1,2	
77	1	*Tylomelania perconica*	190083	5	AY312001 (190083.5)
78	10	*Tylomelania perconica*	190083	9	
		Tylomelania spec. 4	190191	107-111	
		Tylomelania spec. 4	190192	3,5	
		Tylomelania spec. 4	190857	2	KJ850820 (190857.2)
		Tylomelania spec. 4	192134	8	
79	2	*Tylomelania perconica*	190822	1,4	KJ933756 (190822.1)
80	5	*Tylomelania perfecta*	190797	1-3,6	KJ850780 (190797.3)
		Tylomelania wallacei	192182	1	
81	2	*Tylomelania perfecta*	190797	4,5	KJ850781 (190797.4)
82	26	*Tylomelania perfecta*	191639	1	KJ850782 (191639.1)
		Tylomelania spec. 8	192146	1-5	
		Tylomelania robusta	192152	1-6	
		Tylomelania spec. 18	191143	1-8	
		Tylomelania spec. 18	192177	1-6	
83	45	*Tylomelania perfecta*	191639	2	KJ850783 (191639.2)
		Tylomelania spec. 8	190683d	1	
		Tylomelania spec. 8	190688b	1,3	
		Tylomelania spec. 8	190790	1-4,6	
		Tylomelania spec. 8	190793	1-6	
		Tylomelania spec. 8	190810	1,2,4-8	
		Tylomelania spec. 8	191144a	1-4,6	
		Tylomelania spec. 8	191144b	1-6	
		Tylomelania spec. 8	191146	2,3,6	
		Tylomelania robusta	192154	2,6	
		Tylomelania sp.	193900	1,2	
		Tylomelania sp.	193901	1	
		Tylomelania wallacei	190789	1,2	
		Tylomelania wallacei	190792	3	
		Tylomelania wallacei	192164	1	
84	3	*Tylomelania* spec. 4	190191	106	KJ933757 (190191.106)
		Tylomelania spec. 4	192134	7	
		Tylomelania sp.	193902	1	
85	1	*Tylomelania* spec. 4	190192	1	AY312000 (190192.1)
86	4	*Tylomelania* spec. 4	190192	4	
		Tylomelania spec. 4	190844	4-6	KJ850819 (190844.5)
87	1	*Tylomelania* spec. 4	190192	6	KJ933758 (190192.7)
88	1	*Tylomelania* spec. 4	190192	7	KJ933759 (190192.7)
89	48	*Tylomelania* spec. 4	190831	1,3	
		Tylomelania spec. 4	190844	2,3	
		Tylomelania robusta	190811	1	
		Tylomelania spec. 14	190084	1,2,4	AY312017 (190084.2)
		Tylomelania spec. 14	190189	3-8	
		Tylomelania spec. 14	191593a	1-8	
		Tylomelania spec. 14	191593b	1,2,4,5	
		Tylomelania spec. 14	191594	1	
		Tylomelania spec. 14	191596	1-6	
		Tylomelania spec. 14	190851	1,5-8	

Table continued

		Tylomelania spec. 14	190840	1-6	
		Tylomelania spec. 14	190868	1-4	
90	1	*Tylomelania* spec. 4	190831	2	KJ933760 (190831.2)
91	1	*Tylomelania* spec. 4	190831	4	KJ933761 (190831.4)
92	1	*Tylomelania* spec. 4	190831	5	KJ933762 (190831.5)
93	1	*Tylomelania* spec. 4	190831	6	KJ933763 (190831.6)
94	1	*Tylomelania* spec. 4	190831	7	KJ850818 (190831.7)
95	1	*Tylomelania* spec. 4	190831	8	KJ933764 (190831.8)
96	10	*Tylomelania* spec. 4	190857	1,4-6	KJ933765 (190857.1)
		Tylomelania spec. 4	192134	3-6	
		Tylomelania spec. 4	193902	2,4	
97	1	*Tylomelania* spec. 4	190857	3	KJ850821 (190857.3)
98	1	*Tylomelania* spec. 4	190864	1	KJ850822 (190864.1)
99	2	*Tylomelania* spec. 4	190864	2,5	KJ933766 (190864.2)
100	3	*Tylomelania* spec. 4	190864	3,4,6	KJ933767 (190864.3)
101	3	*Tylomelania* spec. 4	190867	1-3	KJ850823 (190867.1)
102	1	*Tylomelania* spec. 4	190867	4	KJ933768 (190867.4)
103	2	*Tylomelania* spec. 4	192134	1	KJ933769 (192134.1)
		Tylomelania spec. 4	193902	3	
104	1	*Tylomelania* spec. 4	192134	2	KJ933770 (192134.2)
105	12	*Tylomelania* spec. 5	190809	1-6	KJ850824 (190809.5)
		Tylomelania spec. 5	190841	1-6	
106	4	*Tylomelania* spec. 5	190850	2,5-7	KJ850826 (190850.5)
107	5	*Tylomelania* spec. 5	191136	1-3	KJ850828 (191136.2)
		Tylomelania spec. 5	191137	2,8	
108	6	*Tylomelania* spec. 5	191137	1,3-7	KJ850829 (191137.1)
109	3	*Tylomelania* spec. 8	115517	1,2	KJ933771 (115517.1)
		Tylomelania wallacei	115516	1	
110	12	*Tylomelania* spec. 8	190683b	2	
		Tylomelania spec. 8	190794	1-3	KJ850843 (190794.3)
		Tylomelania wallacei	190683a	1	
		Tylomelania wallacei	190683f	2	
		Tylomelania wallacei	190685c	2	
		Tylomelania wallacei	190795	4	
		Tylomelania wallacei	192150	6	
		Tylomelania wallacei	192167a	6	
		Tylomelania wallacei	192172a	3,6	
111	9	*Tylomelania* spec. 8	190687	1-3,5,8-12	KJ850833 (190687.2)
112	11	*Tylomelania* spec. 8	190687	4,6,7	
		Tylomelania spec. 8	190689	1-8	KJ850839 (190689.7)
113	1	*Tylomelania* spec. 8	190688a	10	KJ933772 (190688a.10)
114	1	*Tylomelania* spec. 8	190688a	11	KJ933773 (190688a.11)
115	8	*Tylomelania* spec. 8	190688a	1,3,5,7-9	KJ850834 (190688a.3)
		Tylomelania spec. 8	190688b	2,4	
116	3	*Tylomelania* spec. 8	190688a	2,4,6	KJ850835 (190688a.6)
117	4	*Tylomelania* spec. 8	190796	1-4	KJ933774 (190796.1)
118	5	*Tylomelania* spec. 8	190805	1-5	KJ850844 (190805.2)
119	5	*Tylomelania* spec. 8	190806	1,3-6	KJ850845 (190806.1)
120	9	*Tylomelania* spec. 8	190806	2	KJ850846 (190806.2)
		Tylomelania spec. 8	190807	1-4	
		Tylomelania robusta	192155	1-4	
121	21	*Tylomelania* spec. 8	190815a	1-6	KJ850849 (190815a.2)
		Tylomelania spec. 8	190815b	2,4,5,7,8	
		Tylomelania spec. 8	190855	1-4,6	
		Tylomelania spec. 8	191145	1,3-6	
122	4	*Tylomelania* spec. 8	190815b	1,3,6	KJ850852 (190815b.6)
		Tylomelania spec. 8	191145	2	
123	2	*Tylomelania* spec. 8	190855	5	
		Tylomelania spec. 8	190866	5	KJ850858 (190866.5)
124	19	*Tylomelania* spec. 8	190856	1-4	KJ850854 (190856.2)
		Tylomelania spec. 8	190862	1-6	
		Tylomelania robusta	190811	3,4	
		Tylomelania robusta	190845	1-6	
		Tylomelania sp.	193903	1	
125	5	*Tylomelania* spec. 8	190866	1-3,6,8	KJ850857 (190866.1)
126	1	*Tylomelania* spec. 8	190866	4	KJ933775 (190866.4)
127	1	*Tylomelania* spec. 8	190866	7	KJ933776 (190866.7)
128	2	*Tylomelania* spec. 8	190869	1,2	KJ850859 (190869.1)
129	1	*Tylomelania* spec. 8	191144a	5	KJ933777 (191144a.5)
130	3	*Tylomelania* spec. 8	192183	1	KJ933778 (192183.1)
		Tylomelania spec. 8	192188	2	
		Tylomelania wallacei	192182	3	
131	1	*Tylomelania* spec. 8	192188	1	KJ933779 (192188.1)
132	2	*Tylomelania robusta*	190811	2	KJ933780 (190811.2)
		Tylomelania wallacei	190685c	6	

Table continued

133	2	*Tylomelania robusta*	190853	1,4	KJ850786 (190853.4)
134	3	*Tylomelania robusta*	190853	2,3,5	KJ933781 (190853.2)
135	8	*Tylomelania robusta*	190865	1-8	KJ850787 (190865.2)
136	2	*Tylomelania robusta*	192154	1	KJ850789 (192154.1)
		Tylomelania wallacei	192148	1	
137	1	*Tylomelania scalariopsis*	190009	4	AY311962 (190009.4)
138	16	*Tylomelania scalariopsis*	190859	1	KJ850792 (190859.1)
		Tylomelania spec. 20	190832	2	
		Tylomelania spec. 20	191597	1-6	
		Tylomelania spec. 20.	191598	1,3,6-8	
		Tylomelania spec. 20	191599	2,3,7	
139	4	*Tylomelania scalariopsis*	190863	4	KJ850793 (190863.4)
		Tylomelania scalariopsis	190008	3	
		Tylomelania spec. 20	190207	1	
		Tylomelania spec. 20	191605	11	
140	8	*Tylomelania scalariopsis*	191603	1	KJ933782 (191603.1)
		Tylomelania spec. 20	190832	3	
		Tylomelania spec. 20	191598	2,4,5	
		Tylomelania spec. 20	191599	1,5,8	
141	3	*Tylomelania scalariopsis*	191603	2,3	KJ933783 (191603.2)
		Tylomelania spec. 20	191604	2	
142	3	*Tylomelania scalariopsis*	191603	4	KJ933784 (191603.4)
		Tylomelania spec. 20	191604	1,3	
143	1	*Tylomelania* spec. 1	191027a	5	KJ850797 (191027a.5)
144	1	*Tylomelania* spec. 17	190812	2	KJ933785 (190812.2)
145	7	*Tylomelania* spec. 17	190812	3	KJ850870 (190812.3)
		Tylomelania spec. 20	191600	1-6	
146	4	*Tylomelania* spec. 18	191131	1,2	KJ933786 (191131.1)
		Tylomelania spec. 18	191133	7,10	
147	10	*Tylomelania* spec. 18	191131	3	KJ850872 (191131.3)
		Tylomelania spec. 26	195096	14,21	
		Tylomelania wallacei	115515	2,6	
		Tylomelania wallacei	190683a	2	
		Tylomelania wallacei	192185	1,4,6,8	
148	8	*Tylomelania* spec. 18	191132	1-8	KJ933787 (191132.1)
149	8	*Tylomelania* spec. 18	191133	1-6,8,9	KJ850873 (191133.1)
150	31	*Tylomelania* spec. 25	190032	1	KJ933788 (190032.1)
		Tylomelania spec. 25	191113	5,6	
		Tylomelania spec. 25	191612	1,3,6	
		Tylomelania spec. 13	191092a	1-4	
		Tylomelania spec. 13	191092b	2	
		Tylomelania spec. 13	191092c	1,3-5	
		Tylomelania spec. 13	191093	5	
		Tylomelania spec. 13	191112a	3,5,6	
		Tylomelania spec. 13	191112b	1,3,4	
		Tylomelania spec. 13	191118a	2,3,5,6	
		Tylomelania spec. 13	191118b	4,5	
		Tylomelania spec. 13	191613a	1-3	
151	4	*Tylomelania* spec. 25	191113	1,3,4	KJ933789 (191113.1)
		Tylomelania spec. 25	191612	2	
152	1	*Tylomelania* spec. 25	191113	2	KJ933790 (191113.2)
153	5	*Tylomelania* spec. 22	192189	1-4,6	KJ850884 (192189.1)
154	1	*Tylomelania* spec. 22	192189	5	KJ933791 (192189.5)
155	1	*Tylomelania* spec. 13	190035	1	AY312013 (190035.1)
156	4	*Tylomelania* spec. 13	190035	2,3,5,6	KJ933792 (190035.2)
157	1	*Tylomelania* spec. 13	190035	4	KJ933793 (190035.4)
158	6	*Tylomelania* spec. 14	190084	3	KJ933794 (1900843)
		Tylomelania spec. 14	191594	2	
		Tylomelania spec. 14	191595	1-3	
		Tylomelania spec. 14	190851	3	
159	30	*Tylomelania* spec. 15	190187	3,4,6,7	
		Tylomelania spec. 15	190188	3-8	AY312004 (190188.3)
		Tylomelania spec. 15	191138	3,5	
		Tylomelania spec. 15	191139	1-8	
		Tylomelania spec. 15	191140	1-3,8	
		Tylomelania spec. 15	191141a	3,4,7	
		Tylomelania spec. 15	191141b	1,3,6	
160	1	*Tylomelania* spec. 15	190187	5	KJ933795 (190187.5)
161	4	*Tylomelania* spec. 14	190190	1,3,5,7	AY312016 (190190.1)
162	1	*Tylomelania* spec. 14	190190	4	KJ933796 (190190.4)
163	5	*Tylomelania* spec. 16	190206	1	AY312009 (190206.1)
		Tylomelania spec. 20	190846	2,4-6	
164	4	*Tylomelania* spec. 16	190206	2,3,5,6	KJ933797 (190206.2)
165	1	*Tylomelania* spec. 16	190206	4	KJ933798 (190206.4)
166	4	*Tylomelania* spec. 20	190832	1,4,6	KJ933799 (190832.1)

Table continued

		Tylomelania spec. 20	191599	6	
167	1	*Tylomelania* spec. 20	190832	5	KJ850878 (190832.5)
168	1	*Tylomelania* spec. 20	190846	1	KJ933800 (190846.1)
169	1	*Tylomelania* spec. 13	191092c	2	KJ933801 (191092c.2)
170	6	*Tylomelania* spec. 13	191093	2,6	KJ933802 (191093.2)
		Tylomelania spec. 13	191121	1,2,4,5	
171	2	*Tylomelania celebicola*	191094a	1	KJ933803 (191094a.1)
		Tylomelania celebicola	191094b	1	
172	7	*Tylomelania* spec. 13	191112a	1,2,4	KJ933804 (191112a.1)
		Tylomelania spec. 13	191120	1,3	
		Tylomelania spec. 13	191611a	1	
		Tylomelania spec. 13	191613a	4	
173	1	*Tylomelania* spec. 13	191112b	2	KJ933805 (191112b.2)
174	9	*Tylomelania* spec. 13	191114	1-9	KJ933806 (191114.1)
175	9	*Tylomelania* spec. 13	191115	1,3,5,6,10	KJ933807 (191115.1)
		Tylomelania spec. 13	191116a	1,4	
		Tylomelania spec. 13	191116b	2	
		Tylomelania spec. 13	191611b	2	
176	5	*Tylomelania* spec. 13	191115	2,4,7,9,11	KJ933808 (191115.2)
177	1	*Tylomelania* spec. 13	191115	8	KJ933809 (191115.8)
178	2	*Tylomelania* spec. 13	191120	4	KJ933810 (191120.4)
		Tylomelania spec. 13	191611a	6	
179	1	*Tylomelania* spec. 13	191120	5	KJ933811 (191120.5)
180	1	*Tylomelania* spec. 13	191120	7	KJ933812 (191120.7)
181	4	*Tylomelania* spec. 15	191138	1,2,6,8	KJ933813 (191138.1)
182	2	*Tylomelania* spec. 15	191138	4,7	KJ933814 (191138.4)
183	3	*Tylomelania* spec. 15	191140	4	KJ933815 (191140.4)
		Tylomelania spec. 15	191141a	1	
		Tylomelania spec. 15	191141b	2	
184	7	*Tylomelania* spec. 15	191140	5	KJ933816 (191140.5)
		Tylomelania spec. 15	191141a	2,5,6,8	
		Tylomelania spec. 15	191141b	4,5	
185	1	*Tylomelania* spec. 15	191140	6	KJ933817 (191140.6)
186	1	*Tylomelania* spec. 15	191140	7	KJ933818 (191140.7)
187	8	*Tylomelania* spec. 12	191142	1-8	KJ850865 (191142.8)
188	2	*Tylomelania* spec. 12	191143	1,4	KJ933819 (191143.1)
189	7	*Tylomelania* spec. 12	191143	2,3,5-9	KJ933820 (191143.2)
190	1	*Tylomelania* spec. 14	191593b	3	KJ933821 (191593b.3)
191	1	*Tylomelania* spec. 14	191595	4	KJ933822 (191595.4)
192	1	*Tylomelania* spec. 20	191599	4	KJ933823 (191599.4)
193	5	*Tylomelania* spec. 20	191601	1,2	KJ933824 (191601.1)
		Tylomelania spec. 13	192160	4,6,8	
194	1	*Tylomelania* spec. 20	191601	3	KJ933825 (191601.3)
195	9	*Tylomelania* spec. 20	191602	1,3-10	KJ933826 (191602.1)
196	1	*Tylomelania* spec. 20	191602	2	KJ933827 (191602.2)
197	1	*Tylomelania* spec. 20	191605	10	KJ933828 (191605.10)
198	1	*Tylomelania* spec. 20	191605	12	KJ933829 (191605.12)
199	3	*Tylomelania* spec. 20	191605	1,8,9	KJ933830 (191605.1)
200	1	*Tylomelania* spec. 20	191605	2	KJ933831 (191605.2)
201	2	*Tylomelania* spec. 20	191605	3,5	KJ933832 (191605.3)
202	1	*Tylomelania* spec. 20	191605	4	KJ933833 (191605.4)
203	1	*Tylomelania* spec. 20	191605	6	KJ933834 (191605.6)
204	1	*Tylomelania* spec. 20	191605	7	KJ933835 (191605.7)
205	1	*Tylomelania* spec. 13	192158	1	KJ933836 (192158.1)
206	2	*Tylomelania* spec. 13	192158	2,3	KJ933837 (192158.2)
207	2	*Tylomelania* spec. 13	192158	4,6	KJ933838 (192158.4)
208	3	*Tylomelania* spec. 13	192158	5,7,8	KJ933839 (192158.5)
209	4	*Tylomelania* spec. 13	192160	1-3,7	KJ933840 (192160.1)
210	1	*Tylomelania* spec. 13	192160	5	KJ933841 (192160.5)
211	1	*Tylomelania* sp. 4	193899	3	KJ933842 (193899.3)
212	2	*Tylomelania* spec. 14	190851	2,4	KJ933843 (190851.2)
213	2	*Tylomelania* spec. 19	190693	1,2	KJ850877 (190693.1)
214	3	*Tylomelania* spec. 19	190694	1	EF140865 (190694.1)
		Tylomelania spec. 19	190839	2,3	
215	3	*Tylomelania wallacei*	115515	1,3,5	KJ933844 (115515.1)
216	1	*Tylomelania wallacei*	115515	4	KJ933845 (115515.4)
217	1	*Tylomelania wallacei*	115516	2	KJ933846 (115516.2)
218	4	*Tylomelania wallacei*	115516	3	KJ933847 (115516.3)
		Tylomelania wallacei	190683a	4	
		Tylomelania wallacei	190683c	3	
		Tylomelania wallacei	190683e	1	
219	9	*Tylomelania wallacei*	190211	3	
		Tylomelania wallacei	190798a	2,4,5	KJ850902 (190798a.2)
		Tylomelania wallacei	190798b	1-5	
220	1	*Tylomelania wallacei*	190683a	3	KJ933848 (190683a.3)

Table continued

221	4	*Tylomelania wallacei*	190683c	1	KJ933849 (190683c.1)
		Tylomelania wallacei	190795	2	
		Tylomelania wallacei	192172a	5	
		Tylomelania wallacei	192172b	2	
222	1	*Tylomelania wallacei*	190683c	2	KJ933850 (190683c.2)
223	4	*Tylomelania wallacei*	190683e	2	KJ933851 (190683e.2)
		Tylomelania wallacei	190685a	6	
		Tylomelania wallacei	190686	2	
		Tylomelania wallacei	192163	3	
224	11	*Tylomelania wallacei*	190685a	1,5	KJ850893 (190685a.1)
		Tylomelania wallacei	190685b	1	
		Tylomelania wallacei	190685c	5	
		Tylomelania wallacei	190686	1	
		Tylomelania wallacei	192150	4,7	
		Tylomelania wallacei	192163	1	
		Tylomelania wallacei	192167a	4	
		Tylomelania wallacei	192172a	2	
		Tylomelania wallacei	192172b	3	
225	1	*Tylomelania wallacei*	190685a	3	KJ933852 (190685a.3)
226	1	*Tylomelania wallacei*	190792	2	KJ933853 (190792.2)
227	1	*Tylomelania wallacei*	190798b	6	KJ933854 (190798b.6)
228	1	*Tylomelania wallacei*	192167a	1	KJ933855 (192167a.1)
229	9	*Tylomelania wallacei*	192167a	2,5	KJ933856 (192167a.2)
		Tylomelania wallacei	192167b	1-4	
		Tylomelania wallacei	192172a	1,4	
		Tylomelania wallacei	192172b	4	
230	1	*Tylomelania wallacei*	192167a	3	KJ933857 (192167a.3)
231	1	*Tylomelania wallacei*	192182	2	KJ933858 (192182.2)
232	1	*Tylomelania wallacei*	192185	2	KJ933859 (192185.2)
233	3	*Tylomelania wallacei*	192185	3,5,7	KJ933860 (192185.3)

Acknowledgements

First, I thank Thomas von Rintelen (Museum für Naturkunde) for introducing me to Indonesia and Sulawesi, for his confidence and letting me work independently on different research projects and for his continuous financial support during my time at the Museum. Furthermore, I thank the (former) members of the Rintelen lab (Renita Danabalan, Daniela Franz, Kristin Leinhoß, Romy Richter, Kristina von Rintelen and Andreas Wessel) and several colleagues from the Museum for fruitful discussions on snails and systematics and for the broad support when help, a listening ear or a drink was needed after a stressful day.

I thank the Research Center for Biology, Indonesian Institute of Sciences (LIPI) – especially Ristiyanti M. Marwoto, Renny K. Hadiaty and Daisy Wowor – and Kementerian Riset dan Teknologi Republik Indonesia (RISTEK) for the permit to conduct research in Indonesia in recent years, and PT. INCO/PT. VALE Indonesia Tbk. for logistic support in Sulawesi (Malili Lakes). Moreover, I thank Christian Albrecht, Catharina Clewing, Mark de Bruyn, Matthias Glaubrecht, Robert Hall, Fabian Herder, Arne Nolte, Jobst Pfaender, Romy Richter, Mark Schultz, John Slapcinsky, Thomas von Rintelen, and Andreas Wessel for a wonderful and inspiring time in the field, and particularly, for relaxed discussions on biogeography, insects, snails, fishes, projects and local food and beer. Terima kasih!

I am indebted to the Research Department of the Museum für Naturkunde and the molecular lab staff for the financial support to do fieldwork in Indonesia, to analyze data in the lab and to visit various national and international conferences to present my research.

Some of the analyses performed here would have been impossible without the following people: Ladislav Bocak, Sebastian Klaus, David Lohman, Marie-Lilith Patou, Hsi-Te Shih and T.N.C. Vidya provided data necessary for the ancestral state reconstruction analyses in Chapter 1. Jacob A. Esselstyn and Ronald Clouse provided sequence data for phylogenetic analyses in Chapter 2. Johannes Graf and Robert Shelly provided extensive samples of rainbowfish species and several atheriniform samples needed in Chapter 3. Kristin Leinhoß performed and illustrated the haplotype network analyses shown in Chapter 4.

Finally, and most important, I thank my family – particularly my mother!!! – and friends for always supporting me and never losing the faith in me. Last but not least: I am very happy that I attended the SIAL VI conference in Bogor (Indonesia) in 2012 and thus was able to meet Catharina (for the second time). Huge thanks for getting through the hard times of commuting week after week between Gießen and Berlin since that time and for staying patient until we finally moved in couple of weeks ago!

Conference contributions

Stelbrink, B., de Bruyn, M., Albrecht, C., Hall, R. & von Rintelen, T.: "Crossing lines – Biogeographic patterns in Sulawesi" (oral). **BioDivEvo 2014**. 7-11 July 2014, Honolulu, Hawaii, USA.

Stelbrink, B., de Bruyn, M., Albrecht, C., Hall, R. & von Rintelen, T.: "Sulawesi – A biogeographic melting pot" (oral). **BioDivEvo 2014**. 24-27 March 2014, Dresden, Germany.

Stelbrink, B., Richter, R., Glaubrecht, M., Albrecht, C., Hall, R. & von Rintelen, T.: "Where West meets East – The biogeography of Sulawesi revisited" (keynote). **SAGE, 2nd Southeast Asian Gateway Evolution Meeting**. 11-15 March 2013, Berlin, Germany.

de Bruyn, M., **Stelbrink, B.** et al.: "Terrestrial and limnic insights into Southeast Asian biogeography" (keynote). **SAGE, 2nd Southeast Asian Gateway Evolution Meeting**. 11-15 March 2013, Berlin, Germany.

Franz, D., **Stelbrink, B.**, Richter, K., Albrecht, C. & von Rintelen, T.: "New insights into the biogeography of SE Asian freshwater snails (Gastropoda: Lymnaeidae)" (poster). **SAGE, 2nd Southeast Asian Gateway Evolution Meeting**. 11-15 March 2013, Berlin, Germany.

Wessel, A., Mühlethaler, R., von Rintelen, K., von Rintelen, T., **Stelbrink, B.**, Wachmann, E. & Hoch, H.: "The first root community in Southeast Asian caves: Cave planthoppers from Maros karst, Sulawesi (Hemiptera: Fulgoromorpha: Cixiidae: Bennini)" (poster). **SAGE, 2nd Southeast Asian Gateway Evolution Meeting**. 11-15 March 2013, Berlin, Germany.

Gauffre Autelin, P., Bößneck, U., Clewing, C., Stelbrink, B., von Rintelen, T. & Albrecht, C.: "New discoveries in old lakes: Pea-clams (Bivalvia: Sphaeriidae: *Pisidium* spp.) of the Malili lake system (Indonesia, Sulawesi)" (poster). **SAGE, 2nd Southeast Asian Gateway Evolution Meeting**. 11-15 March 2013, Berlin, Germany.

Stelbrink, B. & von Rintelen, T.: "Radula genes and adaptive radiation in the freshwater snail *Tylomelania* in Lake Towuti (Sulawesi)" (poster). **SIAL 6. Speciation in Ancient Lakes 6: Classic concepts & new approaches**. 26-31 August 2012, Bogor, Indonesia.

Danabalan, R., **Stelbrink, B.** & von Rintelen, T.: "Advances in using wet museum collections of molluscs for molecular phylogenetics" (poster). **8th Biennial Conference of the Systematics Association**. 4-8 July 2011, Belfast, Northern Ireland.

Stelbrink, B. & von Rintelen, T.: "The biogeography of Sulawesi - is there evidence for a vicariant origin of taxa on the island?" (oral). **Annual Conference of the Society for Tropical Ecology (GTÖ): Status and Future of Tropical Biodiversity**. 21-24 February 2011, Frankfurt/Main, Germany.

Kriwet, J., **Stelbrink, B** & Tanaka, S.: "Early ontogeny of the frilled shark, Chlamydoselachus anguineus, from Suruga Bay, Japan" (poster). **11th Annual Science Conference of the European Elasmobranch Association (EEA).** 23-26 November 2007, Brest, France.

Freyhof, J. & **Stelbrink, B**.: "Red List of European Cobitidae: Listing of most threatened species for conservation priority" (oral). **3rd International Conference: Loaches of the Genus Cobitis and Related Genera, Biology, Systematics, Genetics, Distribution, Ecology, Conservation.** 24-29 September 2006, Sibenik, Croatia.

Freyhof, J., **Stelbrink, B**., Özulug, M. & Economidis. P.S.: "The first record of Cobitis puncticulata from Europe with comments on its conservation status (Teleostei: Cobitidae)" (poster). **3rd International Conference: Loaches of the Genus Cobitis and Related Genera, Biology, Systematics, Genetics, Distribution, Ecology, Conservation.** 24-29 September 2006, Sibenik, Croatia.

Stelbrink, B. & Freyhof, J.: "Reduktion von Schuppen und Kopfkanälen bei Pomatoschistus canestrinii (Ninni, 1883) (Teleostei: Gobiidae)" (poster). **5. Tagung der Gesellschaft für Ichthyologie (GfI)**. 22-24 September 2005, Bonn, Germany.

List of publications

de Bruyn M.*, **Stelbrink, B.***, Morley R.J., Hall R., Carvalho G.R., Cannon C.H., van den Bergh G., Meijaard E., Metcalfe I., Boitani L., Maiorano L., Shoup R. & von Rintelen T. (2014) Borneo and Indochina are major evolutionary hotspots for Southeast Asian biodiversity. *Systematic Biology,* in press.

*Co-first authors

von Rintelen T., **Stelbrink B.**, Marwoto, R.M. & Glaubrecht (2014) A snail perspective on the biogeography of Sulawesi, Indonesia: origin and intra-island dispersal of the viviparous freshwater gastropod *Tylomelania*. *PLoS ONE*, **9**, e98917.

Stelbrink B., Stöger I., Hadiaty, R.K., Schliewen, U.K. & Herder, F. (2014) Age estimates for an adaptive lake fish radiation, its mitochondrial introgression, and an unexpected sister group: sailfin silversides of the Malili Lakes system in Sulawesi. *BMC Evolutionary Biology*, **14**, 94.

Hoch H., Wessel A., Asche M., Baum D., Beckmann F., Bräunig P., Ehrig K., Mühlethaler R., Riesemeier H., Staude A., **Stelbrink B.**, Wachmann E., Weintraub P., Wipfler B., Wolff C. & Zilch M. (2014b) Non-sexual abdominal appendages in adult insects challenge a 300 million year old bauplan. *Current Biology*, **24**, R16–R17.

Hoch H., Bourgoin T., **Stelbrink B.** & Wessel A. (2014a) Small giants from Madagascan caves: autapomorphic giantism in the new cave-dwelling planthopper *Tsingya clarkei* gen. nov., sp. nov. (Hemiptera: Fulgoroidea: Meenoplidae). *Journal of Natural History*, **48**, 919–935.

de Bruyn M., Rüber L., Nylinder S., **Stelbrink B.**, Lovejoy N.R., Lavoué S., Tan H.H., Nugroho E., Wowor D., Ng P.K.L., Siti Azizah M.N., von Rintelen T., Hall R. & Carvalho G.R. (2013a) Paleo-drainage basin connectivity predicts evolutionary relationships across three Southeast Asian biodiversity hotspots. *Systematic Biology*, **62**, 398–410.

de Bruyn M., **Stelbrink B.**, Page T.J., Phillips M.J., Lohman D.J., Albrecht C., Hall R., von Rintelen K., Ng P.K.L., Shih H.-T., Carvalho G.R. & von Rintelen T. (2013b) Time and space in biogeography: response to Parenti & Ebach (2013). *Journal of Biogeography*, **40**, 2204–2206.

Wessel A., Hoch H., Asche M., von Rintelen T., **Stelbrink B.**, Heck V., Stone F.D. & Howarth F.G. (2013) Founder effects initiated rapid species radiation in Hawaiian cave planthoppers. *Proceedings of the National Academy of Sciences USA*, **110**, 9391–9396.

Günther R., **Stelbrink B.** & von Rintelen T. (2012) Three new species of *Callulops* (Anura: Microhylidae) from western New Guinea. *Vertebrate Zoology*, **62**, 407–423.

Hoch H., Mühlethaler R., Wachmann E., **Stelbrink B.** & Wessel A. (2011) *Celebenna thomarosa* gen. n., sp. n. (Hemiptera, Fulgoromorpha, Cixiidae, Bennini) from Indonesia: Sulawesi with notes on its ecology and behaviour. *Deutsche Entomologische Zeitschrift*, **58**, 241–250.

Stelbrink B., Albrecht C., Hall R. & von Rintelen T. (2012) The biogeography of Sulawesi revisited: is there evidence for a vicariant origin of taxa on Wallace's "anomalous island"? *Evolution*, **66**, 2252–2271.

von Rintelen K., Page T.J., Cai Y., Roe K., **Stelbrink B.**, Kuhajda B.R., Iliffe T.M., Hughes J. & von Rintelen T. (2012) Drawn to the dark side: a molecular phylogeny of freshwater shrimps (Crustacea: Decapoda: Caridea: Atyidae) reveals frequent cave invasions and challenges current taxonomic hypotheses. *Molecular Phylogenetics and Evolution*, **63**, 82–96.

Stelbrink B., von Rintelen T., Cliff G. & Kriwet J. (2010) Molecular systematics and global phylogeography of angel sharks (genus *Squatina*). *Molecular Phylogenetics and Evolution*, **54**, 395–404.

Kriwet J., Endo H. & **Stelbrink B.** (2010) On the occurrence of the Taiwan angel shark, *Squatina formosa* Shen & Ting, 1972 (Chondrichthyes, Squatinidae) from Japan. *Zoosystematics and Evolution*, **86**, 117–124.

Günther R., **Stelbrink B.** & von Rintelen T. (2010) *Oninia senglaubi*, another new genus and species of frog (Amphibia, Anura, Microhylidae) from New Guinea. *Zoosystematics and Evolution*, **86**, 245–256.

Stelbrink B. & Wessel A. (2008) D.E.Z. – a history in numbers. *Deutsche Entomologische Zeitschrift*, **55**, 7–12.

Freyhof J., **Stelbrink B.**, Özulug M. & Economidis P.S. (2008) The first record of *Cobitis puncticulata* from Europe with comments on its conservation status (Teleostei: Cobitidae). *Folia Zoologica*, **57**, 16–19.

Freyhof J. & **Stelbrink B.** (2007) *Cobitis illyrica*, a new species of loach from Adriatic basin, Croatia (Teleostei: Cobitidae). *Ichthyological Exploration of Freshwaters*, **18**, 269–275.

Stelbrink B. & Freyhof J. (2006) Reduction of scales and head canals in *Pomatoschistus canestrinii* (Ninni, 1883) (Teleostei: Gobiidae). *Verhandlungen der Gesellschaft für Ichthyologie*, **5**, 71–77.

Selbständigkeitserklärung

Hiermit erkläre ich, dass ich die vorliegende Arbeit selbständig und nur unter Verwendung der angegebenen Literatur und Hilfsmittel angefertigt habe.

Berlin, Björn Stelbrink

Parts of this thesis have been already published or are currently "in press" in peer-reviewed journals.

de Bruyn M.*, **Stelbrink, B.***, Morley R.J., Hall R., Carvalho G.R., Cannon C.H., van den Bergh G., Meijaard E., Metcalfe I., Boitani L., Maiorano L., Shoup R. & von Rintelen T. (2014) Borneo and Indochina are major evolutionary hotspots for Southeast Asian biodiversity. *Systematic Biology*, **63**, 879-901.

*Co-first authors

B.S., M.d.B and T.v.R. conceived the study; B.S. collected and analyzed data for the molecular meta-analyses, performed molecular clock analyses and ancestral state estimations and visualized the results (Fig. 1.3-5); M.d.B. and B.S. took the leading role in writing the manuscript.

Stelbrink B., Albrecht C., Hall R. & von Rintelen T. (2012) The biogeography of Sulawesi revisited: is there evidence for a vicariant origin of taxa on Wallace's "anomalous island"? *Evolution*, **66**, 2252–2271.

B.S. and T.v.R. conceived the study; B.S. collected and analyzed data for the molecular meta-analyses, illustrated Sulawesi's palaeogeographic history, performed molecular clock analyses and visualized the results (all figures); B.S. and T.v.R. wrote the paper.

Stelbrink B., Stöger I., Hadiaty, R.K., Schliewen, U.K. & Herder, F. (2014) Age estimates for an adaptive lake fish radiation, its mitochondrial introgression, and an unexpected sister group: sailfin silversides of the Malili Lakes system in Sulawesi. *BMC Evolutionary Biology*, **14**, 94.

B.S., U.K.S. and F.H. and conceived the study; B.S. performed the molecular (clock) analyses, visualized the results (all figures), B.S., U.K.S. and F.H. wrote the paper.

von Rintelen T., **Stelbrink B.**, Marwoto, R.M. & Glaubrecht (2014) A snail perspective on the biogeography of Sulawesi, Indonesia: origin and intra-island dispersal of the viviparous freshwater gastropod *Tylomelania*. *PLoS ONE*, **9**, e98917.

T.v.R. and M.G. conceived the study; B.S. performed DNA lab work, analyzed and interpreted molecular phylogenies and haplotype network analyses, performed molecular clock analyses and discussed the manuscript; T.v.R. wrote the paper.